

| | |
|-------------|--|
| Title | Studies on structure-function relationships of soybean glycinin at subunit levels and improvement of its functions by protein engineering(Dissertation_全文) |
| Author(s) | Prak, Krisna |
| Citation | Kyoto University (京都大学) |
| Issue Date | 2006-03-23 |
| URL | http://dx.doi.org/10.14989/doctor.k12368 |
| Right | |
| Type | Thesis or Dissertation |
| Textversion | author |

**Studies on Structure–Function Relationships of Soybean Glycinin at
Subunit Levels and Improvement of Its Functions by Protein
Engineering**

Krisna Prak

2006

Contents

| | Page |
|--|------|
| List of Abbreviations | |
| General Introduction | 3 |
| Chapter I: | |
| Structure–Function Relationships of Soybean Glycinin at Subunit Levels Assessed Using Mutant Lines | 9 |
| Chapter II: | |
| Structure–Function Relationships of Soybean Proglycinins at Subunit Levels | 22 |
| Chapter III: | |
| Protein Engineering of Soybean Proglycinin C-Terminus for Improving Physicochemical Properties | 42 |
| Chapter IV: | |
| Design of Genetically Modified Soybean Proglycinin A1aB1b with Multiple Copies of Bioactive Peptide Sequences | 66 |
| Summary | 85 |
| Acknowledgements | 90 |
| List of Publications | 92 |

List of Abbreviations

| | |
|-----------------|---|
| ANS | 8-anilino-1-naphthalene sulfonic acid |
| <i>p</i> -APMSF | (<i>p</i> -amidinophenyl) methanesulfonyl fluoride |
| BSA | bovine serum albumin |
| DSC | differential scanning calorimetry |
| <i>E. coli</i> | <i>Escherichia coli</i> |
| EDTA | ethylene diaminetetraacetate |
| HPLC | high performance liquid chromatography |
| IPTG | isopropyl-1-thio- β -D-galactopyranoside |
| LB | Luria-Bertani |
| ODS | octadecyl silica |
| PAGE | polyacrylamide gel electrophoresis |
| PCR | polymerase chain reaction |
| PDB | protein data bank |
| SDS | sodium dodecyl sulfate |
| TB | terrific broth |
| WT | wild type |

General Introduction

Soybean is the world's foremost provider of protein and oil. Soybean seed contains 38% protein, 18% oil, 15% soluble carbohydrate, 15% insoluble carbohydrate, and 14% moisture/ash/other. Furthermore, it is rich in calcium, iron, zinc, phosphate, magnesium and B vitamins. Due to its high nutritional value, cheaper price compared to animal meats and milks, and its abundant bioavailability, soybean seed has become a much more attractive raw material for food industries. Soybean foods are available as meat analogs, miso, tempeh, soymilk, soy yogurt, shoyu, tofu, and natto. Besides these, soybean semi-products from the manufacturers have been available in the form of defatted soy flour, enzyme active soy flour, full fat soy, grits, hydrolyzed soy protein, soy fibers, soy protein concentrate, soy protein isolate, textured soy flour, and textured vegetable protein [1].

Soybean proteins have been reported for their many desirable functional properties in food systems. Some of them include gelation, emulsification, foaming, cohesion-adhesion, elasticity, viscosity, solubility, water/fat absorption, and flavor binding [2]. Recently, it has been reported that soybean proteins were used for cheese, various kinds of drinks, and salami production [3]. Aside from physicochemical properties, soybean proteins have been reported for various health benefits such as cholesterol lowering effect [4–9], blood pressure lowering effect [10,11], reduction of plasma triglyceride levels [12,13], appetite suppressing effect [14], antiobesity effect [15], antiatherosclerosis effect [16], and anticarcinogenesis effect [17–19].

Soybean (*Glycine max* L.) proteins are composed of two major components, glycinin (11S globulin) and β -conglycinin (7S globulin), accounting for 40 and 30% of the total seed proteins, respectively [20,21]. β -Conglycinin is a trimeric protein composed of three kinds of subunit: α (~67 kDa), α' (~71 kDa), and β (~50 kDa). Together, these subunits form seven heterotrimers with random combinations as well as three homotrimers. The core regions are highly homologous among the three subunits. In addition, the α and α' subunits have N-terminal extension regions composed of 125 and 141 amino acid residues, respectively, which are rich in acidic amino acid residues. Another difference between the

three subunits is that two carbohydrate moieties are attached to the α and α' subunits, whereas only one is attached to the β subunit. The physicochemical properties of recombinant wild type β -conglycinin homotrimers and of the native α and α' subunits, deletion mutants (α_c and α'_c) without the N-terminal extension regions and mutant β -conglycinin heterotrimers have previously been investigated [22–27]. These studies demonstrated that (i) thermal stability differs between the three subunits in the following order: β (90.8 °C) > α (82.7°C) > α' (78.6°C); (ii) the homologous core regions are responsible for these differences in thermal stability; (iii) the solubility of these subunits depends on the number of N-terminal extensions and carbohydrate moieties; (iv) the surface hydrophobicity is known to be higher in α and α' subunit than the β subunit and their surface hydrophobicities are conferred by the core region of each subunit; (v) the emulsifying abilities and heat-induced associations of the heterotrimers containing one β subunit are similar to those of the α or α' homotrimer, whereas those of the heterotrimers containing two β subunits are similar to those of the β homotrimer; and, (vi) the difference in gel hardness of β -conglycinin does not relate to the secondary structure change after heating, but correlates to the difference in strand thickness and the density of the gel network. Furthermore, physiological function of β -conglycinin has also been improved by introduction of bioactive peptides [28–31].

On the other hand, glycinin is a hexameric protein composed of five kinds of major subunit (A1aB1b, A1bB2, A2B1a, A3B4, and A5A4B3), each of which consists of an acidic (~30 kDa) and a basic (~20 kDa) polypeptide linked by a single disulfide bond, except A4 of A5A4B3 [32]. The five subunits have been classified into two groups based on sequence homology. Group I consists of A1aB1b (~53.6 kDa), A1bB2 (~52.2 kDa), and A2B1a (~52.4 kDa), and group II consists of A3B4 (~55.4 kDa) and A5A4B3 (~61.2 kDa). The homology within a group is more than 84%, and 45–49% between groups [33,34]. So far, only gel forming ability of glycinin was well investigated [35–41]. Therefore, in order to enhance the use of soybean, it is very important to study structure-function relationships of soybean glycinin at subunit level and to improve its functionality by protein engineering. For these reasons, the author at first investigated structure-function relationships of soybean glycinin using mutant soybean lines containing

glycinins composed of individual groups (I and II), A3B4 only, A5A4B3 only, and glycinin containing all five kinds of subunit (Chapter I). To gain insight into structure-function relationship of all glycinin subunits, homogeneous preparation composed of only one kind of molecular species is necessary. However, it is very difficult to obtain a large amount of such homogeneous species from soybean seeds because of their molecular heterogeneity. The author then studied structure-function relationship of soybean glycinin using proglycinin subunits which were obtained by constructing *Escherichia coli* expression system harboring cDNAs of individual subunits (Chapter II). Next, to improve physicochemical functions of glycinin, the author introduced extension regions of β -conglycinin α' and α subunits to proglycinins A1aB1b and A2B1a C-termini, and twenty negatively/positively charged amino acid residues and A5A4B3 hypervariable region to A1aB1b C-terminus. A successful result for the improvement of emulsion property of proglycinins is described in Chapter III. Lastly, the author improved the physiological function of soybean proglycinin by introducing of IIAEK peptide sequences which have a hypocholesterolemic activity greater than β -sitosterol, a drug for hypercholesterolemia, to five disordered regions of proglycinin A1aB1b. Folding ability of mutants and purification of IIAEK peptide from their tryptic hydrolysates were analyzed (Chapter IV).

References

- [1] American Soybean Association, URL http://www.asa-europe.org/SoyInfo/composition_e.htm, 2005 (accessed December 20, 2005).
- [2] Utsumi, S., Matsumura, Y. and Mori, T. (1997). Structure-function relationships of soy proteins. In food proteins and their applications. S. Damodaran and A. Paraf, eds. Marcel Dekker, Inc. New York, NY, pp. 257–291.
- [3] Friedman, M and Brandon, D. L. (2001). Nutritional and health benefits of soy proteins. *J. Agric. Food Chem.* **49**, 1069–1086.
- [4] Carroll, K. K. (1982). Hypercholesterolemia and atherosclerosis: effects of dietary protein. *Fed. Proc.* **41**, 2792–2796.
- [5] Anderson, J. W., Johnstone, B. M. and Cook-Newell, M. E. (1995). Meta-analysis of the effects of soy protein intake on serum lipids. *N. Engl. J. Med.* **333**, 276–282.
- [6] Potter, S. M. (1995). Overview of proposed mechanisms for the hypocholesterolemic effect of soy. *J. Nutr.* **125**, 606S–611S.
- [7] Fukui, K., Tachibana, N., Wanezaki, S., Tsuzaki, S., Takamatsu, K., Yamamoto, T., Hashimoto, Y. and Shimoda, T. (2002). Isoflavone-free soy protein prepared by column chromatography reduces plasma cholesterol in rats. *J. Agric. Food Chem.* **50**, 5717–5721.
- [8] Bakhit, R. M., Klein, B. P., Essex-Sorlie, D., Ham, J. O., Erdman, J. W. and Potter, S. M. (1994). Intake of 25g of soybean protein with or without soybean fiber alters plasma lipids in men with elevated cholesterol concentrations. *J. Nutr.* **124**, 213–222.
- [9] Huang, W., Wood, C., L'Abbé, Gilani G. S., Cockell K. A. and Xiao, C. W. (2005). Soy protein isolate increases hepatic thyroid hormone receptor content and inhibits its binding to target genes in rats. *J. Nutr.*, **135**, 1631–1635.
- [10] Lo, E. M. Y. and Li-Chan, E. C. Y. (2005). Angiotensin I converting enzyme inhibitory peptides from in vitro pepsin-pancreatin digestion of soy protein. *J. Agric. Food Chem.* **53**, 3369–3376.
- [11] Ringseis, R., Matthes B., Lehmann, V., Becker, K., Schöps R., Ulbrich-Hoffmann, R. and Eder, K. (2005). Peptides and hydrolysates from casein and soy protein modulate the release of vasoactive substances from human aortic endothelial cells. *Biochim. Biophys. Acta.* **1721**, 89–97
- [12] Iritani, N., Hosomi, H., Fukuda, H., Tada, K. and Ikeda, H. (1996). Soybean protein suppresses hepatic lipogenic enzyme gene expression in wistar fatty rats. *J. Nutr.* **126**, 380–388.
- [13] Sugano, M., Tanaka, K. and Ide, Y. (1982). Secretion of cholesterol, triglyceride and apolipoprotein A-I by isolated perfused liver from rats fed soybean protein and casein or their amino acid mixtures. *J. Nutr.* **112**, 855–862.
- [14] Nishi, T., Hara, H., Asano, K. and Tomita, F. (2003). The soybean β -conglycinin β 51–63 fragment suppresses appetite by stimulating cholecystokinin release in rats. *J. Nutr.* **133**, 2537–2542.

- [15] Aoyama, T., Fukui, K., Takamatsu, K., Hashimoto, Y. and Yamamoto, T. (2000). Soy protein isolate and its hydrolysate reduce body fat of dietary obese rats and genetically obese mice (Yellow KK). *Nutrition* **16**, 349–354.
- [16] Huff, M. W., Roberts, C. K. D. and Carroll, K. K. (1982). Long-term effects of semipurified diets containing casein or soy protein isolate on atherosclerosis and plasma lipoprotein in rabbits. *Atherosclerosis* **41**, 327–336.
- [17] Messina, M. J., Persky, V., Setchell, K. D. R. and Barnes, S. (1994). Soy intake and cancer risk. *Nutr. Cancer* **21**, 113–131.
- [18] Adlercreutz, C. H. T., Goldin, B. R., Gorbach, S. L., Hockerstedt, K. A. V., Watanabe, S., Hamalainen, E. K., Markkanen, M. H., Makela, T. H., Wahala, K. T., Hase, T. A. and Fotsis, T. (1995). Soybean phytoestrogen intake and cancer risk. *J. Nutr.* **125**, 757S–770S.
- [19] Tsuruki, T., Takahata, K. and Yoshikawa, M. (2005). Anti-alopecia mechanisms of soymetide-4, an immunostimulating peptide derived from soy β -conglycinin. *Peptides* **26**, 707–711.
- [20] Utsumi, S. (1992). Plant food protein engineering. In *Advances in Food and Nutrition Research*; Kinsella, J. E., Ed.; Academic Press: SanDiego, CA, pp 89–208.
- [21] Utsumi, S., Matsumura, Y and Mori, T. (1997). Structure-function relationships of soy proteins. In *Food Proteins and Their Applications*; Damodaran, S., Paraf, A., Eds.; Marcel Dekker: New York, pp 257–291.
- [22] Maruyama, N., Mohamad Ramlan, M. S., Takahashi, K., Yagasaki, K., Goto, H., Hontani, N., Nakagawa, S. and Utsumi, S. (2002). The effect of the N-linked glycans on structural features and physicochemical functions of soybean β -conglycinin homotrimers. *J. Am. Oil Chem. Soc.* **79**, 139–144.
- [23] Maruyama, N., Mohamed Salleh, M. R., Tkahashi, K., Yagasaki, K., Goto, H., Hontani, N., Nakagawa, S. and Utsumi, S. (2002). Structure-physicochemical function relationships of soybean β -conglycinin heterotrimers. *J. Agric. Food Chem.* **50**, 4323–4326.
- [24] Maruyama, N., Katsube, T., Wada, Y., Oh, M. H., Barba de la Rosa, A. P., Okuda, E., Nakagawa, S. and Utsumi, S. (1998). The roles of the N-linked glycans and extension regions of soybean β -conglycinin in folding, assembly and structural features. *Eur. J. Biochem.* **258**, 854–862.
- [25] Maruyama, N., Sato, R., Wada, Y., Matsumura, Y., Goto, H., Okuda, E., Nakagawa, S. and Utsumi, S. (1999). Structure-physicochemical function relationships of soybean β -conglycinin constituent subunits. *J. Agric. Food Chem.* **47**, 5279–5284.
- [26] Salleh, M. R. B. M., Maruyama, N., Takahashi K., Yagasaki, K., Higasa, T., Matsumura Y. and Utsumi, S. (2004). Gelling properties of soybean β -conglycinin having different subunit compositions. *Biosci. Biotechnol. Biochem.* **68**, 1091–1096.
- [27] Takenaka, Y., Doyama, N, Maruyama, N., Utsumi, S. and Yoshikawa, M. (2004). Introduction of DPR, an enterostatin fragment peptide, into soybean β -conglycinin α' subunit by site-directed mutagenesis. *Biosci. Biotechnol. Biochem.* **68**, 253–256.

- [28] Matoba, N., Doyama, N., Yamada, Y., Maruyama, N., Utsumi, S. and Yoshikawa, M. (2001). Design and production of genetically modified soybean protein with anti-hypertensive activity by incorporating potent analogue of ovokinin (2-7). *FEBS lett.* **497**, 50–54.
- [29] Maruyama, Y. (2004). Crystal structure of soybean β -conglycinin and its application to introduction of bioactive peptides. *PhD thesis*, Laboratory of food design and development, faculty of agriculture, Kyoto university. Kyoto, Japan.
- [30] Onishi, K., Matoba, N., Yamada, Y., Doyama, N., Maruyama, N., Utsumi, S., Yoshikawa, M. (2004). Optimal designing of β -conglycinin to genetically incorporate RPLKPW, a potent anti-hypertensive peptide. *Peptides* **25**, 37–43.
- [31] Maruyama, N., Maruyama, Y., Tsuruki, T., Okuda, E., Yoshikawa, M., Utsumi, S. (2003). Creation of soybean β -conglycinin β with strong phagocytosis-stimulating activity. *Biochim. Biophys. Acta* **1648**, 99–104.
- [32] Staswick, P. E., Hermodson, M. A. and Nielsen, N. C. (1984). The amino acid sequence of the A2B1a subunit of glycinin. *J. Biol. Chem.* **259**, 13424–13430.
- [33] Utsumi, S., Matsumura, Y. and Mori, T. (1997). Structure-function relationships of soy proteins. In *Food Proteins and Their Applications*; Damodaran, S., Paraf, A., Eds.; Marcel Dekker: New York, pp 257–291.
- [34] Nielsen, N. C., Dickinson, C. D., Cho, T. J., Thanh, V. H., Scallon, B. J., Fischer, R. L., Sims, T. L., Drews, G. N. and Goldberg, R. B. (1989). Characterization of the glycinin gene family in soybean. *Plant Cell* **1**, 313–328.
- [35] Staswick, P. E., Hermodson, M. A. and Nielsen, N. C. (1984). The amino acid sequence of the A2B1a subunit of glycinin. *J. Biol. Chem.* **259**, 13424–13430.
- [36] Nielsen, N. C., Dickinson, C. D., Cho, T. J., Thanh, V. H., Scallon, B. J., Fischer, R. L., Sims, T. L., Drews, G. N. and Goldberg, R. B. (1989). Characterization of the glycinin gene family in soybean. *Plant Cell* **1**, 313–328.
- [37] Damodaran, S. (1988). Refolding of thermally unfolded soy proteins during the cooling regime of the gelation process: effect on gelation. *J. Agric. Food Chem.* **36**, 262–269.
- [38] Nakamura, T., Utsumi, S. and Mori, T. (1984). Network structure formation in thermally induced gelation of glycinin. *J. Agric. Food Chem.* **32**, 349–352.
- [39] Babajimopoulos, M., Damodaran, S., Rizvi, S. H. and Kinsella, J. E. (1983). Effects of various anions on the rheological and gelling behavior of soy proteins: thermodynamic observations, *J. Agric. Food Chem.* **31**, 1270–1275.
- [40] Chanyongvorakul, Y., Matsumura, Y., Sakamoto, Y., Motoki, M. and Mori, T. (1994). Gelation of beans 11S globulins by Ca^{2+} -independent transglutaminase. *Biosci. Biotech. Biochem.* **58**, 864–869.
- [41] Kang, I. J., Matsumura, Y., Ikura, K., Motoki, M., Sakamoto, H. and Mori, T. (1994). Gelation and gel properties of soybean glycinin in a transglutaminase-catalyzed system. *J. Agric. Food Chem.* **42**, 159–165.

Chapter I

Structure–Function Relationships of Soybean Glycinin at Subunit Levels Assessed Using Mutant Lines

Many molecular species that have different subunit compositions with random combinations are present in soybean seeds [1]. From normal soybean cultivars, it is very difficult to isolate large quantities of glycinin molecular species of restricted subunit compositions, which are necessary for the investigation of structure-physicochemical function relationships of glycinin at the subunit level. Zhang et al. [2] succeeded in the isolation of glycinin molecular species composed of only group I subunits using a soybean cultivar lacking A5A4B3. Nakamura et al. [3,4] prepared artificial glycinins with a homogeneous subunit composition by renaturation from isolated subunits [4] or polypeptides [3]. However, these methods are tedious and laborious. Recently, mutant soybean lines containing glycinin composed of only group I, only group II, only A3B4, or only A5A4B3 were developed [5,6]. In this study, the author purified such glycinins from these mutant soybean lines and compared their physicochemical properties with those of glycinin containing all five kinds of subunit.

Materials and Methods

Purification of Glycinins with Various Subunit Compositions

Glycinins composed of all subunits (11S), only group I subunits (group I-glycinin), only group II subunits (group II-glycinin), only A3B4 (A3B4-glycinin), and only A5A4B3 (A5A4B3-glycinin) were partially purified from corresponding normal and mutant soybean lines [5,6] by the procedure of Nagano et al. [7]. The partially purified fractions containing glycinins were subjected to ammonium sulfate fractionation: 11S, 50–65% saturation; group I-glycinin, 50–75%; group II-glycinin, 50–70%; A3B4-glycinin, 50–65%; and A5A4B3-glycinin, 50–70%. Their purities were assessed by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) [8].

Protein Measurement

The protein concentration of the samples was determined using a Protein Assay Rapid Kit (Wako, Osaka, Japan) based on dye binding with bovine serum albumin as the standard.

Thermal Stability

The thermal stability of the glycinin samples in buffer A [35 mM sodium phosphate, pH 7.6, 0.4 M NaCl, 10 mM 2-mercaptoethanol, 1 mM ethylenediaminetetraacetic acid, 0.1 mM (*p*-amidinophenyl) methanesulfonyl fluoride, and 0.02% NaN₃] was analyzed using a Microcal MC-2 ultrasensitive microcalorimeter (Micro Cal Inc., Northampton, MA) as described previously [9]. All analyses were performed with a protein concentration of 0.5 mg/ml.

Surface Hydrophobicity

Surface hydrophobicities of the glycinin samples were measured by hydrophobic column chromatography using Phenyl Sepharose 6 Fast Flow and Butyl Sepharose 4 Fast Flow columns (both from Amersham Pharmacia Biotech.) as described previously [10]. The glycinin samples were dialyzed against buffer A containing 2.3 M ammonium sulfate and then applied to columns equilibrated with the same buffer. The adsorbed samples were eluted with a linear gradient (2.3–0 M) of ammonium sulfate over a period of 80 min at a flow rate of 0.25 ml/min.

Solubility as a Function of pH

Solubilities of the glycinin samples were measured as described previously [11]. The protein solutions (0.8 mg/ml) were kept at 4 °C for 18 h at various pH values at $\mu = 0.5$ after which the protein solutions were centrifuged, and the protein concentrations in the supernatant were determined as described in the section on protein measurement. The solubility was expressed as a percentage of the total protein content in the sample.

Emulsifying Ability

The emulsifying abilities of the glycinin samples were measured as described previously [11]. To prepare emulsions, 0.25 ml of soybean oil and 1.5 ml of the glycinin samples in buffer A (0.5 mg/ml) were homogenized for 30 s with a high-speed homogenizer (model NS-50, Nichion Irikakikai Ltd.) and further sonicated using an ultrasonic homogenizer (model US-150, Nihonseiki Kaisha Ltd.). The particle size distribution of the emulsions was measured using a laser diffraction instrument (model LA500, Horiba Seisakusho Ltd.). Each sample was measured several times, and a representative typical pattern was presented.

Results and Discussion

Isolation of 11S, Groups I-, Group II-, A3B4- and A5A4B3-Glycinins

Glycinins composed of all five subunits (11S), only group I subunits (group I-glycinin), only group II (group II-glycinin), only A3B4 (A3B4-glycinin), and only A5A4B3

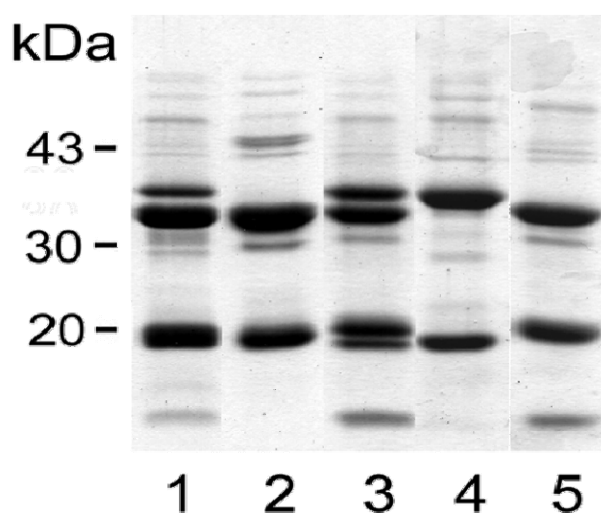


Figure 1. SDS-PAGE analysis of the purified glycinin samples. The purified glycinin samples were analyzed by means of SDS-PAGE using 11% gels. Lane 1, 11S; lane 2, group I-glycinin; lane 3, group II-glycinin; lane 4, A3B4-glycinin; and lane 5, A5A4B3-glycinin. The numbers on the left denote molecular masses.

(A5A4B3-glycinin) were partially purified by the method of Nagano et al. [7] and further purified by ammonium sulfate fractionation. The isolated and purified glycinins had purities of around 90% as evaluated by SDS–PAGE (Figure 1) and were used for the following analyses.

Thermal Stability

The thermal stability of proteins is an important factor for their heat-induced association and gelation. Partial denaturation of the native protein molecule is a prerequisite to the subsequent association of denatured molecules and the formation of the network structure [12,13]. Differential scanning calorimetry (DSC) profiles of the glycinin samples are shown in Figure 2.

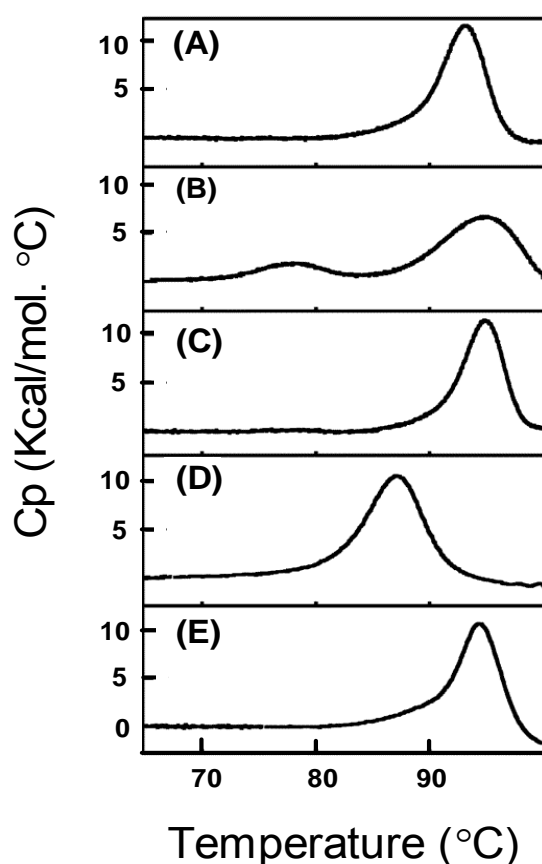


Figure 2. DSC scans of glycinin samples. (A) 11S, (B) group I-glycinin, (C) group II-glycinin, (D) A3B4-glycinin, and (E) A5A4B3-glycinin.

The thermal denaturation midpoint temperature (T_m) of the peaks of the glycinin samples ranged from 77 to 95 °C. 11S containing all five subunits gave a sharp peak with a T_m of 92.9 °C. The T_m of A3B4 was 87.2 °C, which was the lowest among all glycinin species. Even 11S and group II-glycinin, which contain A3B4, exhibited higher T_m values than did A3B4-glycinin. Previously, it was demonstrated that the order of T_m values of β -conglycinin subunits, α , α' , and β , is $\alpha < \alpha' < \beta$, and T_m values of β -conglycinin composed of two or three kinds of subunits were basically conferred by the subunit that had the lowest T_m value [10,14]. Therefore, this is not the case with glycinin.

Group I-glycinin gave two broad peaks (Figure 2B). Their T_m values were 77.0 and 95.0 °C. Utsumi's group developed transgenic tobacco [15] and rice [16] accumulating glycinin A1aB1b and demonstrated that the expressed A1aB1b can be processed and can assemble into hexamers. However, the assembly efficiency was less than 50% and most remained as trimers. Recently, it was further demonstrated using an *Escherichia coli* expression system that A1aB1b had a higher folding ability than A3B4 and assisted in the folding of A3B4 and its modified versions, which had a very low folding ability [17]. Moreover, collaboration between Takaiwa's and Utsumi's groups revealed that A3B4 assists in the incorporation of A1aB1b into hexamers (this will be described elsewhere). These facts suggest that the peaks with T_m values of 95.0 and 77.0 °C originated from a hexamer and a trimer, respectively. This suggestion is consistent with the observation by Utsumi et al. [18] that a glycinin trimer (a half molecule of a hexamer) exhibits a much lower thermal stability than a glycinin hexamer. The fact that the 11S containing both groups I and II subunits exhibit one sharp peak (Figure 2A) supports the above suggestion. On the other hand, both peaks that group I-glycinin gave (Figure 2B) were broader than those of the other glycinin samples (Figure 2A, C–E), suggesting that group II subunits contribute to the structural rigidity of glycinin hexamers.

Tezuka et al. reported that the T_m values of group I-, A3B4-, and A5A4B3-glycinins are 95, 97.9, and 96 °C, respectively [19]. These are not in agreement with the author's results probably because of the different measuring conditions used in the analysis of T_m values.

Surface Hydrophobicity

Surface hydrophobicity of proteins is related to some of their physicochemical properties such as emulsifying and forming abilities and solubility [20,21]. Kato and Nakai [22] described that *cis*-parinaric acid (CPA) is more suitable for the measurement of surface hydrophobicity than 1-anilino-8-naphthalenesulfonate (ANS). Hayakawa and Nakai [23] reported that hydrophobicity measured by phenyl sepharose column chromatography and ANS correlated well to the protein insolubility, whereas no significant correlation was observed between that by CPA and insolubility. They suggested that the aromatic hydrophobicity may play a more important role in protein solubility than the aliphatic hydrophobicity. Therefore, the author employed two columns of phenyl and butyl sepharose (Table 1). With this analysis, the longer the elution time is, the higher the surface hydrophobicity will be of the sample.

Table 1. Elution times of glycinin samples on hydrophobic chromatography

| columns | elution times (min) | | | | |
|------------------|-------------------------------------|----------|-----------------|-------------------|---------------|
| | group I-glycinin | 11S | A5A4B3-glycinin | group II-glycinin | A3B4-glycinin |
| butyl sepharose | 39.9 ^a /0.2 ^b | 43.4/0.3 | 43.9/0.3 | 48.3/0.2 | 52.5/0.2 |
| phenyl sepharose | 57.7/0.3 | 60.9/0.3 | 61.2/0.2 | 66.6/0.2 | 71.7/0.3 |

^a The value represents an average for three experiments. ^b Standard error

Both columns gave similar trends, with A3B4-glycinin having the highest hydrophobicity, and the order was A3B4-glycinin > groupII-glycinin > A5A4B3-glycinin > 11S > group I-glycinin, indicating that there should not be any significant difference in aromatic and aliphatic hydrophobicities among glycinin samples. The proportions of groups I and II subunits in 11S and that of A3B4 and A5A4B3 in group II-glycinin were around 3:2 and 1:1, respectively. This indicates that the surface hydrophobicity of a heterohexamer is consistent with the arithmetic mean of those of its constituent subunits.

Previously, Utsumi's group measured the surface hydrophobicities of the homotrimers of the constituent subunits α , α' , and β by the same procedure [10] and reported that the

order of the surface hydrophobicity is $\alpha > \alpha' > \beta$. From the comparison of the data obtained here and previously, the surface hydrophobicity of β is lower than that of group I-glycinin and that of α is intermediate between those of A3B4 and group II-glycinins. Recently, Tezuka et al. measured the surface hydrophobicities of group I-glycinin, A3B4-glycinin, A5A4B3-glycinin, and β -conglycinin using ANS and reported that A5A4B3-glycinin > A3B4-glycinin > group I-glycinin > β -conglycinin [19]. The discrepancies in the results between Tezuka's and the author's are probably due to the difference in the measuring method used by the two laboratories.

Solubility as a Function of pH

The solubility of proteins is the most important factor for their physicochemical properties [23,24]. Because the author's group previously measured the solubilities of glycinin samples at $\mu = 0.08$ [25], the author measured them at a higher μ of 0.5 (Figure 3). All glycinin samples exhibited a lower solubility at a lower pH with different extents. In particular, group II and A5A4B3-glycinins exhibited low solubilities below pH 5 and their profiles were very similar to each other. A3B4-glycinin was completely soluble at pH 4.8, but 40% soluble at pH 4.0, and less than 20% below pH 3.5. 11S and group I-glycinin became gradually insoluble with a lowering of pH, and their profiles were about the same. These results indicate that at $\mu = 0.5$, the solubilities of 11S and group II-glycinin are governed by those of group I subunits and A5A4B3, respectively. The acidic amino acid contents of group I subunits, A3B4, and A5A4B3 are 56 (average), 66, and 85%, respectively. Therefore, the higher the content of the acidic amino acids, the lower the solubility of the glycinin samples at low pH.

Hayakawa and Nakai [23] reported that the aromatic hydrophobicity of proteins correlates with their insolubility. However, the present results indicate that there is no such relationship between them. One possible reason is that the conditions for measurement of solubility are different. They measured the solubility of proteins in 10 mM phosphate buffer (pH7.0). This condition is closer to author's low ionic strength condition ($\mu = 0.08$) than the high ionic strength condition. Under $\mu = 0.08$ and pH 7.0, the order of solubility was reported to be A5A4B3-glycinin = group II-glycinin > A3B4-glycinin = 11S > group

I-glycinin [25]. There is no correlation between this order and the order of surface hydrophobicity that the author obtained, indicating that the suggestion by Hayakawa and Nakai [23] is not applicable to glycinins.

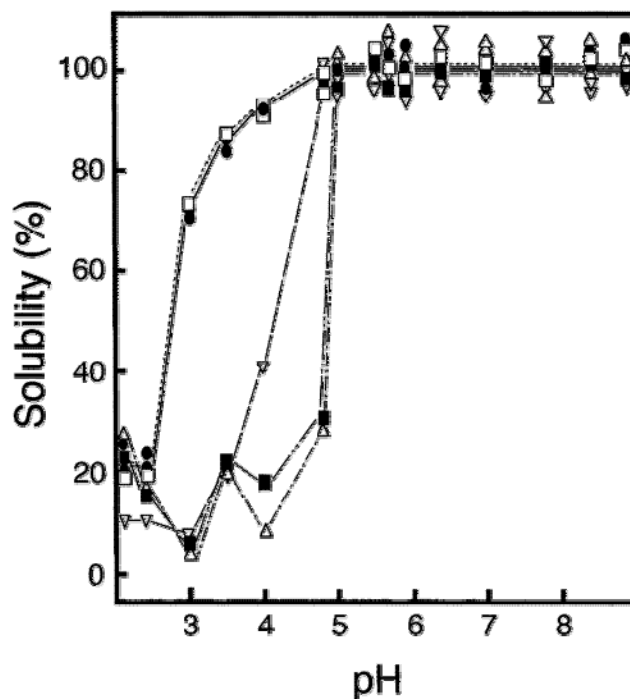


Figure 3. Dependency of the solubility of glycinin samples on pH at ionic strength 0.5. 11S, group I-glycinin, group II-glycinin, A3B4-glycinin, and A5A4B3-glycinin are shown by a solid line with closed circles, a dotted line with open squares, a dashed and double-dotted line with closed squares, a dashed line with inverted triangles, and a dashed and single-dotted line with open triangles, respectively.

All examined molecular species with different subunit compositions of β -conglycinin are completely soluble at pH 3–9 and $\mu = 0.5$ [10,11,14], although all examined glycinin samples exhibited a lower solubility at pH < 5. Glycinin and β -conglycinin are derived from a common ancestor and share a common three-dimensional structure [26–29]. Therefore, differences in the distribution of hydrophobic and charged residues on molecular surfaces may be related to the solubilities of these glycinins.

Emulsifying Ability

The emulsifying ability of proteins is one of the most important physicochemical properties for food processing. The author assessed the emulsifying abilities of glycinin samples by measuring the sizes of their emulsions (Figure 4). 11S and group I-glycinin gave broad peaks with average particle sizes of 3.7 and 6.8 μm , respectively. Group II-, A3B4-, and A5A4B3-glycinins gave sharp peaks with average particle sizes of 2.6, 3.1, and 2.3 μm , respectively. The emulsifying ability of group II-glycinin was better than that of group I-glycinin, and that of 11S was intermediate between those of groups I- and II-

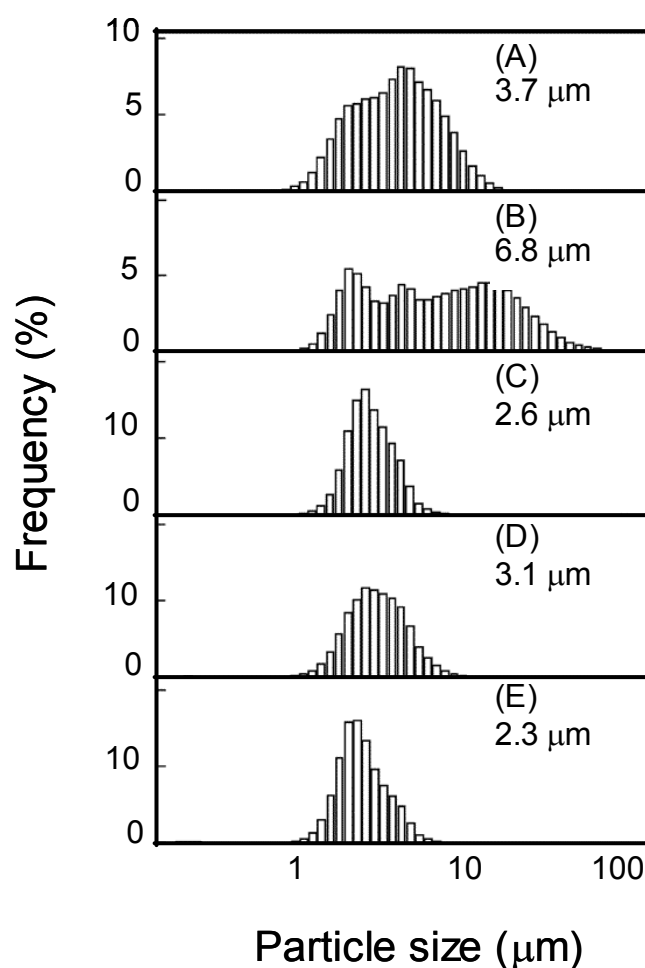


Figure 4. Particle size distributions of emulsions from glycinin samples: (A) 11S, (B) group I-glycinin, (C) group II-glycinin, (D) A3B4-glycinin, and (E) A5A4B3-glycinin.

glycinins. Similarly, the emulsifying ability of A5A4B3-glycinin was better than that of A3B4-glycinin, and that of group II-glycinin was intermediate between those of A5A4B3- and A3B4-glycinins. These results suggest that the emulsifying ability of a heterohexamer corresponds to the arithmetic mean of the emulsifying ability values of its constituent subunits.

Group I-glycinin composed of three different subunits gave very broad peaks. This suggests that the three subunits exhibit a different emulsifying ability to each other. This is consistent with the observation that 11S-containing group I subunits also gave a broad peak. However, the broadness of the 11S peak was much milder than that of the group I-glycinin peak. In analogy with the case of thermal stability, group II subunits are likely to contribute to the sharpness of 11S peak.

Suitable conformational change of proteins at the interface between oil and water and their balance of hydrophilicity and hydrophobicity are important factors for emulsifying ability [30]. Similarly, correlations of surface properties including emulsifying activity with conformational stabilities [31] and surface hydrophobicities of proteins [22] are reported. Previously, it was demonstrated that the structural stability as well as the extension region are important in the emulsifying abilities of β -conglycinin homotrimers [11]. However, the results obtained from the analyses of glycinin samples indicate that there is no relationship between emulsifying ability and structural stability assessed by DSC (Figure 2) and surface hydrophobicity (Table 1) as measured by hydrophobic chromatography. This discrepancy may be due to the difference in the methods employed for the assessment, or such relationships may not be the same for glycinin.

Each subunit of glycinin has several disordered regions, which are not visible by X-ray crystallography [82,29]. The longest one of them is called a hypervariable region [32], which is rich in acidic amino acid residues. The lengths of the hypervariable regions are different among subunits, thus A5A4B3 (103) > A3B4 (70) > group I (A1aB1b, 42; A1bB2, 29; and A2B1a, 35). This order is consistent with that of emulsifying ability. Therefore, it seems that there is a relationship between emulsifying ability and the length

of the hypervariable region of each subunit. This relationship is similar to the important role that the extension region plays in the emulsifying property of β -conglycinin [11].

Comparison of some physicochemical properties of glycinin samples that have restricted subunit compositions from mutant soybean lines with those of normal glycinin indicated that individual subunits have intrinsic properties and that glycinin composed of only group I subunits is suitable for foods requiring high solubility at acidic pH and high ionic strength and glycinin composed of only group II subunits for food requiring emulsifying ability and/or high solubility at neutral pH and low ionic strength.

References

- [1] Utsumi, S. (1992). Plant food protein engineering. In *Advances in Food and Nutrition Research*; Kinsella, J. E., Ed.; Academic Press: SanDiego, CA, pp 89–208.
- [2] Zhang, G. Y., Hayashi, Y., Matsumoto, S., Matsumura, Y. and Mori, T. (2002). Molecular species of glycinin in some soybean cultivars. *Phytochemistry* **60**, 675–681.
- [3] Nakamura, T., Utsumi, S., Kitamura, K., Harada, K. and Mori, T. (1984). Cultivar differences in gelling characteristics of soybean (*Glycine max*) glycinin. *J. Agric. Food Chem.* **32**, 647–651.
- [4] Nakamura, T., Utsumi, S. and Mori, T. (1985). Formation of pseudoglycinins from intermediary subunits of glycinin and their gel properties and network structure. *Agric. Biol. Chem.* **49**, 2733–2740.
- [5] Teraishi, M., Takahashi, M., Hajika, M., Matsunaga, R., Uematsu, Y. and Ishimoto, M. (2001). Suppression of soybean β -conglycinin genes by dominant gene, Scg-1. *Theor. Appl. Genet.* **103**, 1266–1272.
- [6] Takahashi, M., Uematsu, Y., Kashiwaba, K., Yagasaki, K., Hajika, M., Matsunaga, R., Komatsu, K. and Ishimoto, M. (2003). Accumulation of high levels of free amino acids in soybean seeds through integration of mutations conferring seed protein deficiency. *Planta* **217**, 577–586.
- [7] Nagano, T., Hirotsuka, M., Mori, H., Kohyama, K. and Nishinari, K. (1992). Dynamic viscoelastic study on the gelation of 7S globulin from soybean. *J. Agric. Food Chem.* **40**, 941–944.
- [8] Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.
- [9] Maruyama, N., Katsube, T., Wada, Y., Oh, M. H., Barba de la Rosa, A. P., Okuda, E., Nakagawa, S. and Utsumi, S. (1998). The roles of the N-linked glycans and extension regions of soybean β -conglycinin in folding, assembly and structural features. *Eur. J. Biochem.* **258**, 854–862.
- [10] Maruyama, N., Mohamad Ramlan, M. S., Takahashi, K., Yagasaki, K., Goto, H., Hontani, N., Nakagawa, S. and Utsumi, S. (2002). Structure-physicochemical function relationships of soybean β -conglycinin heterotrimers. *J. Agric. Food Chem.* **50**, 4323–4326.
- [11] Maruyama, N., Sato, R., Wada, Y., Matsumura, Y., Goto, H., Okuda, E., Nakagawa, S. and Utsumi, S. (1999). Structure-physicochemical function relationships of soybean β -conglycinin constituent subunits. *J. Agric. Food Chem.* **47**, 5278–5284.
- [12] Utsumi, S., Matsumura, Y. and Mori, T. (1997). Structure-function relationships of soy proteins. In *Food Proteins and Their Applications*; Damodaran, S., Paraf, A., Eds.; Dekker: New York, pp 257–291.
- [13] Clark, A. H., Lea-Tuffnell, C. D. (1986). Gelation of globular proteins. In *Functional Properties of Food Macromolecules*; Mitchell, J. R., Ledword, D. A., Eds.; Elsevier Applied Science: London, pp 203–272.
- [14] Maruyama, N., Mohamad Ramlan, M. S., Takahashi, K., Yagasaki, K., Goto, H., Hontani, N., Nakagawa, S. and Utsumi, S. (2002). The effect of the N-linked glycans on structural features and physicochemical functions of soybean β -conglycinin homotrimers. *J. Am. Oil Chem. Soc.* **79**, 139–144.
- [15] Takaiwa, F., Katsube, T., Kitagawa, S., Higasa, T., Kito, M. and Utsumi, S. (1995). High level accumulation of soybean glycinin in vacuole-derived protein bodies in the endosperm tissue of transgenic tobacco seed. *Plant Sci.* **111**, 39–49.
- [16] Katsube, T., Kurisaka, N., Ogawa, M., Maruyama, N., Ohtsuka, R., Utsumi, S. and Takaiwa, F. (1999). Accumulation of soybean glycinin and its assembly with the glutelins in rice. *Plant Physiol.* **120**, 1063–1073.

- [17] Choi, S. K., Adachi, M., Yoshikawa, M., Maruyama, N. and Utsumi, S. (2004). Soybean glycinin A1aB1b subunit has a molecular chaperone-like function to assist the folding of the other subunit having low folding ability. *Biosci., Biotechnol., Biochem.* **68**, 1991–1994.
- [18] Utsumi, S., Nakamura, T., Harada, H. and Mori, T. (1987). Occurrence of dissociable and undissociable soybean glycinin. *Agric. Biol. Chem.* **51**, 2139–2144.
- [19] Tezuka, M., Yagasaki, K. and Ono, T. (2004). Changes in characters of soybean glycinin groups I, IIa, and IIb caused by heating. *J. Agric. Food Chem.* **52**, 1693–1699.
- [20] Nakai, S. (1983). Structure-function relationships of food proteins with an emphasis on the importance of protein hydrophobicity. *J. Agric. Food Chem.* **31**, 676–683.
- [21] Nakai, S. and Li-Chan, E. (1988). Hydrophobicity-functionality relationship of food proteins. *Hydrophobic Interactions in Food Systems*; CRC Press: Boca Raton, FL, pp 23–41.
- [22] Kato, A. and Nakai, S. (1980). Hydrophobicity determined by a fluorescence probe method and its correlation with surface properties of proteins. *Biochim. Biophys. Acta* **624**, 13–20.
- [23] Hayakawa, S. and Nakai, S. (1985). Relationships of hydrophobicity and net charge to the solubility of milk and soy proteins. *J. Food Sci.* **50**, 486–491.
- [24] Kinsella, J. E. (1979). Functional properties of soy proteins. *J. Am. Oil Chem. Soc.* **56**, 242–258.
- [25] Mori, T., Maruyama, N., Nishizawa, K., Higasa, T., Yagasaki, K., Ishimoto, M. and Utsumi, S. (2004). The composition of newly synthesized proteins in the endoplasmic reticulum determines the transport pathways of soybean seed storage proteins. *Plant J.* **40**, 238–249.
- [26] Maruyama, N., Adachi, M., Takahashi, K., Yagasaki, K., Kohno, M., Takenaka, Y., Okuda, E., Nakagawa, S., Mikami, B. and Utsumi, S. (2001). Crystal structures of recombinant and native soybean β -conglycinin β homotrimers. *Eur. J. Biochem.* **268**, 3595–3604.
- [27] Maruyama, Y., Maruyama, N., Mikami, B. and Utsumi, S. (2004). Structure of the core region of the soybean β -conglycinin α' subunit. *Acta Crystallogr.* **D60**, 289–297.
- [28] Adachi, M., Takenaka, Y., Gidamis, A. B., Mikami, B. and Utsumi, S. (2001). Crystal structure of soybean β -conglycinin A1aB1b homotrimer. *J. Mol. Biol.* **305**, 291–305.
- [29] Adachi, M., Kanamori, J., Masuda, T., Yagasaki, K., Kitamura, K., Mikami, B. and Utsumi, S. (2003). Crystal structure of soybean 11S globulin: Glycinin A3B4 homohexamer. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 7395–7400.
- [30] Damodaran, S. (1997). Protein-stabilized forms and emulsions. In *Food Proteins and Their Applications*; Damodaran, S., Paraf, A., Eds.; Marcel Dekker: New York, pp 57–110.
- [31] Kato, A. and Yutani, K. (1988). Correlation of surface properties with conformational stabilities of wild-type and six mutant tryptophan synthase α -subunits substituted at the same position. *Protein Eng.* **2**, 153–156.
- [32] Argos, P., Narayana, S. V. L. and Nielsen, N. C. (1985). Structural similarity between legumin and vicilin storage proteins from legumes. *EMBO J.* **4**, 1111–1117.

Chapter II

Structure-Function Relationships of Soybean Proglycinins at Subunit Levels

In analogy with Utsumi group's work on β -conglycinin, the author studied the physicochemical properties of glycinin composed of only A3B4, only A5A4B3, only group I, and only group II using mutant soybean cultivars (see Chapter I). Unfortunately, the author was not able to obtain individual molecular species of glycinin composed of only one kind of subunit. By employing the *E. coli* expression system, the author can prepare such individual molecular species, although they are in the form of proglycinin. Utsumi's group has already succeeded in cloning cDNAs encoding A1aB1b [1], A2B1a [2], and A3B4 [3] and also has established the *E. coli* expression systems for A1aB1b [4] and A3B4 [3] subunit precursors of glycinin. In the present study, the author cloned cDNAs for the remaining A1bB2 and A5A4B3 subunits by the reverse transcription-polymerase chain reaction (RT-PCR) method; constructed *E. coli* expression systems for A1bB2, A2B1a, and A5A4B3; and then compared the structural and physicochemical properties of all five subunits.

Materials and Methods

Preparation of cDNAs Encoding Individual Glycinin Subunits and Construction of their Expression Plasmids

Total RNA was extracted from developing soybean seeds according to Shirzadegan et al. [5]. For purification of mRNA from total RNA, the mRNA purification Kit (Amersham Pharmacia Biotech) was used according to the manufacturer's instructions. Full-length cDNAs encoding proglycinin were produced by reverse transcription of mRNA using RNA LA PCR Kit (AMV) Ver. 1.1 (TAKARA). The reaction was performed at 42 °C for 15 min, 60 °C for 15 min, and 99 °C for 5 min.

cDNAs encoding proglycinin subunits A5A4B3 were amplified by PCR using full-length cDNAs as a template for individual subunits. The primers used for the A5A4B3 subunit were 5'-ATT AGC TCC AGC AAG CTC AAC GAG TGC-3' and 5'-C CGC GGA **TCC** **TTA** TGC GAC TTT AAC ACG GGG TGA GC-3' for N and C termini, respectively (underlined and bold letters are *Bam*HI site and a stop codon, respectively). Thirty cycles of PCR were performed using LA TAQ (TAKARA) at 94 °C for 30 s, 60 °C for 30 s, and 72 °C for 4 min.

To avoid misannealing caused by the high homology of the nucleotides sequences of A1bB2 to A1aB1b, which is only one nucleotide difference in 42 nucleotides coding the starting 14 amino acids at the N-terminal of the mature forms, two pairs of primers were used. The first pair served for high selectivity in order to get the cDNA encoding A1bB2 lacking 29 base pairs for the N terminus of the mature form. The second pair of primers was used to get the complete cDNA containing the 29 base pairs based on the sequence reported by Cho et al. [6]. The first pair of primers was 5'-CGA GTGCCA GAT CCA ACG C-3' for the N-terminal and 5'-CTA CGC ACA CTA ACT AGT GC-3' for the C-terminal. The reaction was the same to that of A5A4B3. Thirty cycles of PCR were performed using KOD-Plus (TOYOBO) at 94 °C for 30 s and 68 °C for 2 min and 30 s to amplify the cDNA encoding A1bB2 after LA TAQ (TAKARA). The second pair of primers was 5'-TTC AGT TTC AGA GAG CAG CCA CAG CAA AAC GAG TGC CAG ATC CAA CGC CTC AAT GC-3' and 5'-G CAT GGA **TCC** AGG GCT **CTA** AGC CAC AAC TCT CCT CTG AGA CTC CTT-3' for N and C termini, respectively (*Bam*HI site and a stop codon are indicated as underlined and bold letters, respectively).

The cDNA encoding the proglycinin subunit A2B1a was recovered from pGST4-3-1-4 [2] using KOD-plus and the primers 5'-CTG AGA GAG CAG GCA CAG CAA AAT-3' and 5'-GAG TGC GGC CGC AAG **CTT** **CTA** AGC CAC-3' for the N and C termini, respectively. Underlined and bold letters indicate *Hind*III site and a stop codon, respectively.

The cDNAs encoding the proglycinin subunits A1bB2, A2B1a, and A5A4B3 were blunted using the Blunting Kit (TAKARA) and then digested with *Bam*HI or *Hind*III. The

resulting large DNA fragments were inserted into *Nco*I (filled-in) and *Bam*HI or *Hind*III sites of pET-21d vector (Novagen) to construct the expression plasmids pEA1bB2, pEA2B1a, and pEA5A4B3.

The nucleotide sequences of the cDNAs in the expression plasmids were confirmed by the dideoxy chain termination method of Sanger et al. [7], using the Applied Biosystems sequencer model 310 and the ABI dye terminator cycle sequencing kit with *Amplitaq* polymerase FS (Perkin-Elmer, Applied Biosystems).

Expression of Recombinant Proglycinins

The host cells BL21(DE3), HMS174(DE3), AD494(DE3), and Origami(DE3) containing individual expression plasmids pEA1aB1b [4], pEA1bB2, pEA2B1a, pEA3B4 [3], and pEA5A4B3 were cultured in normal LB and TB media and in media containing additional NaCl. The cells were grown at 37 °C. Expression was induced when A_{600} reached 0.4 to ~0.6 with 1 mM isopropyl-1-thio- β -D-galactoside (IPTG), at 18, 20, and 25 °C. After cultivation, cells were harvested by centrifugation at 5000 rpm for 20 min at 4 °C and stored at –20 °C until used. Aliquots of the cells were analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE) [8] using 11% gel as described previously [9]. Expressed recombinant proglycinins were identified based on their expected sizes and confirmed by western blotting [10] using anti-glycinin antibody followed by goat anti-rabbit IgG-alkaline phosphatase conjugate (Promega).

Purification of Recombinant Proglycinins

All purification steps were carried out at 4 °C and centrifugations were at 6000 rpm for 20 min unless otherwise stated. The basic buffer for all purification steps was buffer A [35 mM potassium phosphate, pH 7.6, 1 mM EDTA, 10 mM 2-mercaptoethanol, 0.1 mM (*p*-amidinophenyl)-methylsulfonyl fluoride, 1 μ g/mL pepstatin A, and 1 μ g/ml leupeptin]. Ammonium sulfate fractionation followed the procedure of Green et al. [11].

Proglycinins A1aB1b, A2B1a, and A3B4. Frozen cells containing proglycinins A1aB1b, A2B1a, and A3B4 were resuspended in buffer B (buffer A containing 0.4 M NaCl) at a

density of 25 g/l buffer and lysed by sonication on an ice bath. Insoluble matter was removed by centrifugation. Expressed proteins were initially fractionated using ammonium sulfate: 35% saturation for A1aB1b and A3B4 and 15% saturation for A2B1a. The precipitate was removed by centrifugation, and the soluble fraction containing recombinant proteins was applied on a Toyopearl (Butyl-650M) (TOSOH) column (2.6 cm × 20 cm) equilibrated with buffer B containing 30% ammonium sulfate for A1aB1b and A3B4 and 15% saturation for A2B1a. Elution was carried out with a linear gradient (800 ml) from 30 to 0% ammonium sulfate in buffer B for A1aB1b and A3B4 and from 15 to 0% for A2B1a. Fractionation of A2B1a was further continued with 100 ml of buffer B. Fractions containing proglycinins A1aB1b, A2B1a, and A3B4 were dialyzed against buffer C (buffer A containing 0.15 M NaCl) and clarified by centrifugation. A1aB1b was applied on Mono Q HR 10/10 column (Pharmacia Biotech) equilibrated with buffer C. Elution was performed with a linear gradient from 0.15 to 0.4 M NaCl in buffer A over a period of 120 min at 2 ml/min. A2B1a and A3B4 were applied on HiLoad 26/10 Q-Sepharose high performance column (Pharmacia Biotech) equilibrated with buffer C. Elution was performed with a linear gradient from 0.15 to 0.5 M NaCl in buffer A over a period of 120 min at 2 ml/min.

Proglycinin A1bB2 and A5A4B3. Frozen cells for proglycinin A1bB2 and A5A4B3 were resuspended in buffer D (buffer A containing 1.0 M NaCl) at a density of 25 g/l buffer and disrupted by sonication on an ice bath. The insoluble matter was removed by centrifugation. Expressed proteins were initially fractionated by 37.5% ammonium sulfate saturation. After removal of precipitates, soluble fractions were further fractionated by ammonium sulfate saturation to 50 and 60% for A1bB2 and A5A4B3, respectively. The precipitates containing A1bB2 and A5A4B3 were then dissolved in buffer D containing 30% ammonium sulfate and then subjected to Toyopearl (Butyl-650M) column (2.6 cm × 20 cm) following the same procedure for A1aB1b and A3B4. Fractions containing A1bB2 and A5A4B3 were dialyzed against buffer D, concentrated by Centriprep YM-10 (Millipore), and then were applied on a gel filtration column (Hi-Prep 26/60 Sephacryl S-300 HR) using buffer D as a mobile phase. Furthermore, A1bB2 and A5A4B3 were subjected to HiLoad 26/10 Q-Sepharose high performance chromatography (Pharmacia Biotech). Elution was performed similar to A2B1a and A3B4 except for the NaCl

concentration of the gradient buffers: 0.15–0.45 M NaCl and 0.25–0.5 M NaCl for A1bB2 and A5A4B3, respectively.

Protein Measurement

The amount of protein in the samples was determined using a Protein Assay Rapid Kit (Wako) with bovine serum albumin as a standard.

Analysis of Self-Assembly into Trimers

Self-assembly of individual proglycinins was analyzed using Hi-Prep 16/60 Sephacryl S-200 HR column as described previously [12].

Solubility as a Function of pH

Proglycinins were dialyzed against buffer E [10 mM sodium phosphate, pH 7.6, 0.5 M NaCl, 1 mM EDTA, 0.1 mM (*p*-amidinophenyl)-methylsulfonyl fluoride, 1 µg/ml pepstatin A, 1 µg/ml leupeptin, 0.02% NaN₃, and 10 mM 2-mercaptoethanol]. Ionic strength and pH were adjusted by adding the appropriate amount of 10 and 50 mM of citrate buffer (pH 2.5–6.5), sodium phosphate buffer (pH 5.5–7.8), and ammonium buffer (pH 7.2–9.0) for the ionic strengths 0.08 and 0.5, respectively. The protein samples (0.8 mg/ml) were kept at 4 °C for 18 h and then centrifuged at 15000 rpm at 4 °C for 15 min to separate the soluble and insoluble fractions. The percent solubility (soluble fraction) was determined by comparing the protein content of the resulting solution with the initial protein content of the sample (100% soluble).

Surface Hydrophobicity

Surface hydrophobicities of proglycinins were analyzed using butyl sepharose 4FF and phenyl sepharose 6FF as described previously [13] except for the starting concentration of ammonium sulfate [35% saturation instead of 2.3 M (48.75%)]. The proteins were eluted with a linear gradient of ammonium sulfate 35–0% over a period of 100 min at a flow rate of 0.25 ml/min.

Differential Scanning Calorimetry (DSC) Measurement

DSC measurement of proglycinins was carried out as described previously [12] using 1 mg/ml of protein. The scan rate was 1 °C/min from 30 to 110 °C under nitrogen gas.

Emulsifying Property

The emulsifying properties of proglycinins were analyzed as described previously [14] using 1.5 ml of 0.5 mg/ml of protein at pH 7.6 and ionic strengths 0.5 and 0.08, mixing with 0.25 ml of soybean oil.

Results and Discussion

Cloning and Sequencing of cDNAs Encoding A1bB2 and A5A4B3 Proglycinins

Because of the high similarity to other group I subunits [15,16] and lower detectable levels of mRNA for proglycinin A1bB2 than those for the other group I subunits in the seed [6], it was difficult to get a cDNA encoding proglycinin A1bB2 by the RT-PCR method using a primer corresponding to its mature N-terminal region. To overcome this problem, two pairs of primers were used (see Materials and Methods). The first pair of primers was of high selectivity to produce a portion of the A1bB2 cDNA lacking 29 base pairs from the first codon for the mature form. A second pair of primers was employed to get the complete cDNA by using the incomplete cDNA as a template. The cDNA for proglycinin A5A4B3 subunit was cloned without any problem.

Nucleotide sequences of the cloned cDNAs were determined and were compared with those in the GenBank protein translation database. Codes used for A1bB2 and A5A4B3 are GMGY3 and AB004062, respectively. All coding sequences determined here were identical to those in the reference database except for one base difference in A1bB2 cDNA; the base A at position 460 from the initiation codon of GMGY3 gene was replaced by C, but there was no change in the amino acid.

Expression and Purification of Individual Proglycinins

To achieve an expression of soluble recombinant proglycinins at a high level, the author attempted cultivation of various host cells having individual expression plasmids under various conditions. In the case of A1bB2, no recombinant protein was found using normal LB medium. However, increasing the salt concentration from 0.17 to 0.5, 0.8, and 1 M resulted in a remarkable expression of soluble recombinant protein (data not shown) with 70, 80, and > 90% solubilities, respectively. On the other hand, NaCl concentration did not affect the A5A4B3 expression level but affected its solubility when using *E. coli* AD494(DE3). Origami(DE3) and BL21(DE3) were the suitable hosts for the expression of A5A4B3 and A3B4, respectively. The most suitable condition for individual proglycinins is summarized in Table 1. In optimizing expression conditions, the author found that it was easier to get soluble expression of group I proglycinins than group II ones, suggesting that the folding ability of group I proglycinin is higher than that of group II.

Table 1. Host cell and culture condition for proglycinin

| proglycinin | host cell (Novagen) | culture conditions for protein expression | | | |
|-------------|------------------------|---|-----------------|--------------|------------------|
| | | medium | salt | temp (°C) | duration (hr) |
| A1aB1b | AD494 (DE3) | LB | 0.17 M NaCl | 25 | 40 |
| A2B1a | AD494 (DE3) | LB | 0.17 M NaCl | 25 | 40 |
| A1bB2 | AD494 (DE3) | LB | 1.0 M NaCl | 25 | 40 |
| A3B4 | BL21 (DE3) | TB | No | 18 | 40 |
| A5A4B3 | Origami (DE3) | TB | 0.1 M Phosphate | 20 | 40 |

The solubility of recombinant proteins from individual expression plasmids in *E. coli* cells was analyzed by SDS-PAGE after sonication (Figure 1). On the basis of band intensity, the solubility of all recombinant proglycinins was estimated to be > 80%. These were confirmed by Western blotting (data not shown) using purified antiglycinin antibody.

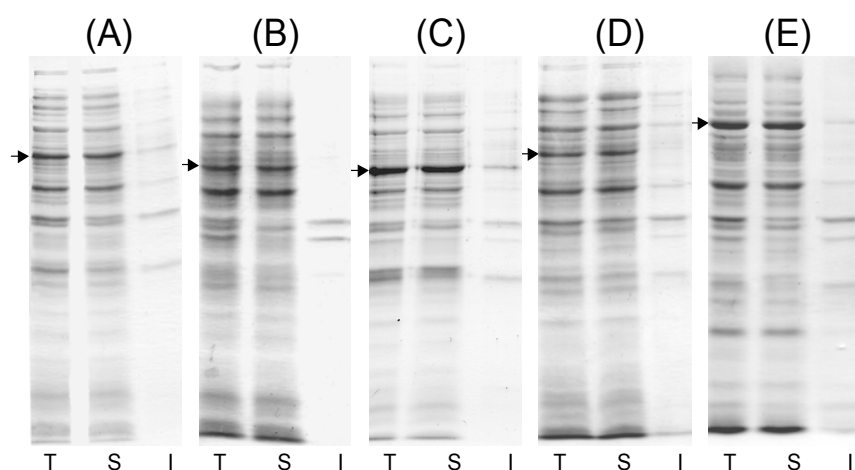


Figure 1. SDS-PAGE (11% acrylamide) profile of the protein expression of recombinant proglycinin subunits. The arrow indicates the position of the expressed proglycinin subunit. T, total protein after cell extraction; S, soluble fraction; and I, insoluble fraction. (A) A1aB1b, (B) A1bB2, (C) A2B1a, (D) A3B4, and (E) A5A4B3.

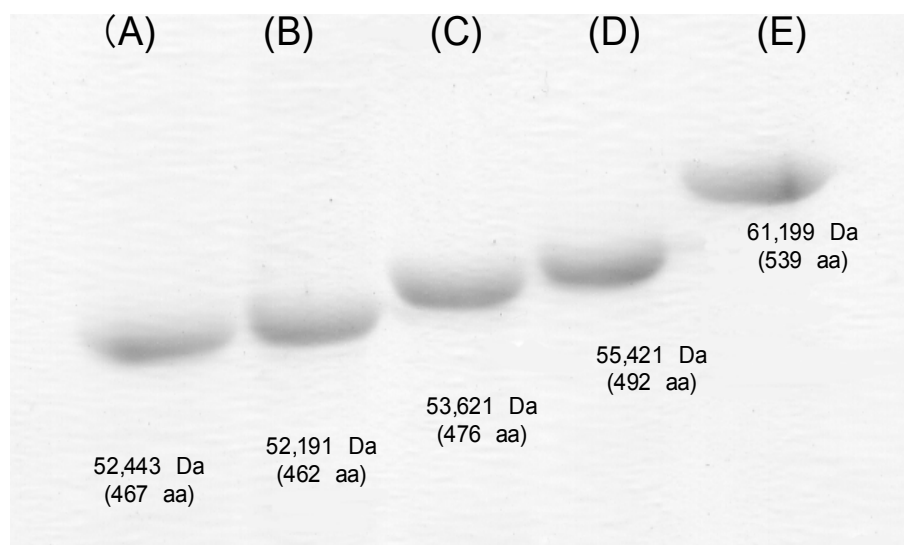


Figure 2. SDS-PAGE analysis of the purified proglycinins. The purified proglycinins were analyzed by SDS-PAGE using 11% acrylamide gel and running at 150 V for 55 min. The number with Da indicates the molecular mass of proglycinin. The number with aa indicates the number of amino acids. (A) A2B1a, (B) A1bB2, (C) A1aB1b, (D) A3B4, and (E) A5A4B3.

After purification, the purity of the proteins was assessed by SDS–PAGE (Figure 2) and found to be > 95%. The number of amino acids and the molecular masses of individual proglycinin subunits calculated based on their cDNA nucleotide sequences are shown in Figure 2. The author found that the mobility of A2B1a was a bit faster than that of A1bB2 although its molecular mass was slightly bigger than A1bB2. Therefore, it must be the mobility of A1bB2 or A2B1a that was inconsistent with its molecular mass. To further analyze the phenomenon of mobility, the author drew a curve Log (molecular mass) of proglycinin subunits vs migration of subunits on the gel (data not shown). The author found that the correlation coefficient of the curve was higher when using A1bB2 with the other three subunits. This means that the mobility of proglycinin subunits was consistent with the number of their amino acids and their molecular masses except those of A2B1a.

Self-Assembly of Proglycinin Subunits

To assess self-assembly into trimers of individual proglycinin subunits, each purified proglycinin sample was subjected to gel filtration chromatography using Hi-Prep 16/60 Sephacryl S-200 HR column at pH 7.6 and $\mu = 0.5$. The elution times were 87.3, 93.2, 95.0, 98.2, and 100.7 min for A5A4B3, A3B4, A1aB1b, A1bB2, and A2B1a, respectively (Figure 3). Previously, it was demonstrated that A1aB1b self-assembles into trimers [9,16,17] and elutes at 95 min on the same column under the same conditions [12]. Elutions of A5A4B3 and A3B4 were faster than that of A1aB1b, and those of A1bB2 and A2B1a were slower. These differences were proportional to their differences in mobilities on SDS–PAGE (Figure 2). Therefore, all proglycinins were assessed to be trimers. Utsumi's group has proposed three criteria for judging the formation of proper conformation of proglycinin: (i) a high level of expression in *E. coli* (> 10% of *E. coli* total proteins), (ii) solubilities comparable with that of globulins, and (iii) self-assembly into trimers [18,19]. Because all of the proglycinins constructed here satisfied these criteria, the author concluded that they assumed a conformation similar to that of native proglycinins. Therefore, *E. coli* expression system in this study could be used for the production of individual proglycinins. The recombinant proteins can then be used to study structural and physicochemical properties of individual proglycinins.

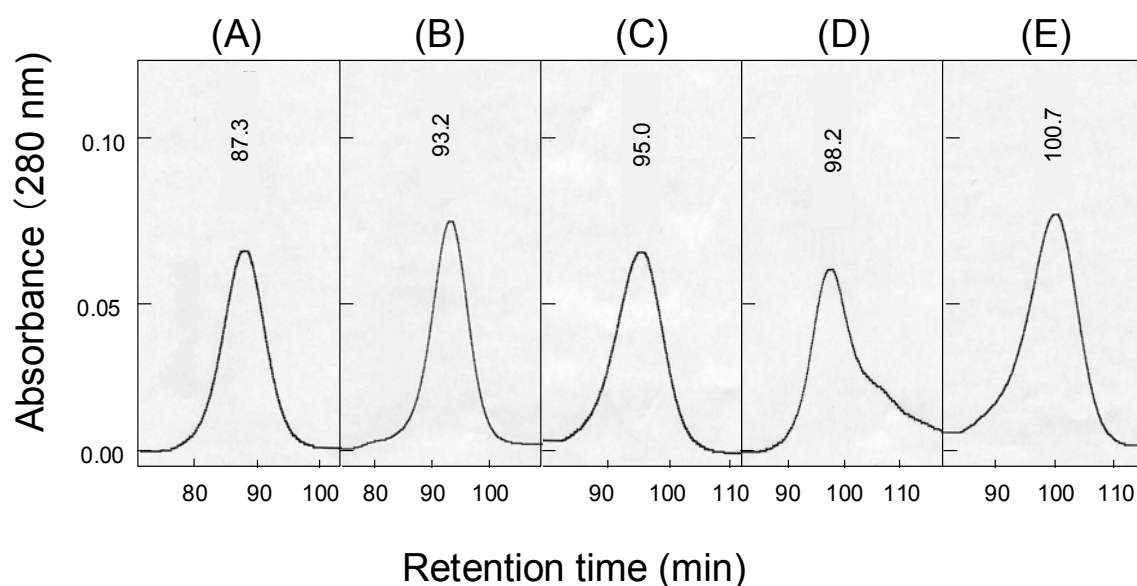


Figure 3. Gel filtration profile of proglycinins. The purified recombinant proglycinins were subjected to chromatography using a Sephacryl S-200 HR column. (A) A5A4B3, (B) A3B4, (C) A1aB1b, (D) A1bB2, and (E) A2B1a.

The elution time of each proglycinin did not follow the order of their real molecular masses but followed that of their mobilities on SDS-PAGE. Because A2B1a has five amino acid residues more than A1bB2 (Figure 2), it should move slightly slower than A1bB2 on SDS-PAGE, but on the contrary, it moved slightly faster. These indicate that A2B1a has the most compact structure among all proglycinin trimers and monomers. Previously, Utsumi's group demonstrated that even one amino acid replacement of glycinin A3 polypeptide results in slower mobility on SDS-PAGE than the normal one [20]. This phenomenon is similar to the unusual behavior of A2B1a on SDS-PAGE. The unusual behavior of A2B1a on gel filtration is a first report for major seed storage proteins. In soybean seeds, many molecular species of different subunit compositions are present [21]. This indicates that individual proglycinins can form heterotrimers. Because they share the same backbone structure, the difference will therefore be on the molecular surface. To understand the difference in the molecular surface of A2B1a subunit, crystallography of A2B1a homotrimer is desired and is in progress by Utsumi's group.

Solubility as a Function of pH

Solubility is a fundamental property for physicochemical functions of proteins for food usage. The author examined the solubility of proglycinin subunits at different pH values at low ($\mu = 0.08$) and high ($\mu = 0.5$) ionic strengths (Figure 4).

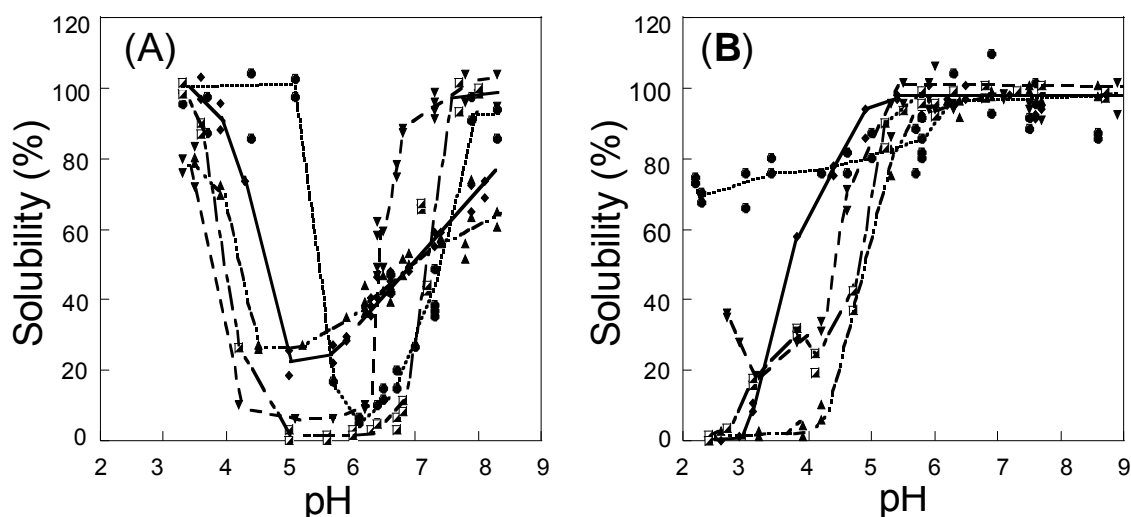


Figure 4. pH dependence of the solubility of proglycinins at ionic strengths of 0.08 (A) and 0.5 (B). A1aB1b, A2B1a, A1bB2, A3B4, and A5A4B3 are shown by a dashed line with closed circles, double-dashed and double-dotted lines with closed triangles, a solid line with closed diamonds, single-dashed and single-dotted lines with half-open squares, and a dashed line with inverted triangles, respectively.

At low ionic strength, A1bB2 and A2B1a exhibited similar solubilities. The lowest solubility was around 20–30% at the pH values 4.5–5.9 and 5.0–5.9 for A2B1a and A1bB2, respectively. Increasing the pH enhanced the solubility of these two subunits but only up to about 60–80%. A1aB1b at the same ionic strength had the narrowest range of insolubility among all proglycinins. In the range of pH 5.7–6.7, its solubility was less than 20%. Group II proglycinins had a wider range of insolubility at $\mu = 0.08$ than group I ones. The solubilities of A3B4 and A5A4B3 were less than 10% at pH 5.0–6.8 and 6–10% at pH 4.2–6.2, respectively. These phenomena were probably due to the fact that A3B4 and A5A4B3 subunits are richer in acidic and basic amino acids than the others.

At high ionic strength, all proglycinins were 100% soluble at $\text{pH} \geq 5.8$. At lower pH, all proglycinins exhibited low solubility except for A1aB1b, which still had good solubility that never went below 60%. A2B1a, A3B4, and A5A4B3 decreased in solubility to nearly 0, 0–25, and 18–36%, respectively. A1bB2 had very low solubility only at $\text{pH} \leq 3.2$.

The pH dependence of the solubility of proglycinins was intrinsic to each subunit. The numbers of amino acid residues, especially charged residues, in the variable regions are different among the five subunits. This difference may have a big effect on the intrinsic solubilities of individual proglycinins. All proglycinins were completely soluble at $\text{pH} \geq 6.0$ at high ionic strength. However, at low ionic strength even at $\text{pH} > 7.6$, the solubilities of A1bB2 and A2B1a were less than 70%. The author cannot explain these phenomena at this point, but they can be studied using mutational approaches and three-dimensional structures.

Utsumi's group isolated glycinin molecular species with restricted subunit compositions: glycinin composed of only group I subunits (group I-glycinin), only group II subunits (group II-glycinin), only A3B4 (A3B4-glycinin), and only A5A4B3 (A5A4B3-glycinin) and studied their solubility at $\mu = 0.08$ [22] and 0.5 (see Chapter I). At $\mu = 0.08$, the solubility profiles of A3B4- and A5A4B3-glycinins were very similar to those of A3B4 and A5A4B3, and that of group I-glycinin was just like a mixture of those of A1aB1b, A1bB2, and A2B1a. On the other hand, at $\mu = 0.5$, the solubility profiles of A3B4- and A5A4B3-glycinins were similar to those of A3B4 and A5A4B3 although there were some differences between pH 4 and pH 5 and that of group I-glycinin was just like a mixture of those of A1aB1b, A1bB2, and A2B1a. These suggest that proglycinin and mature glycinin exhibit similar solubility profiles as a function of pH. These results are consistent with the suggestions based on the solubility of various glycinin molecular species that the solubilities of newly synthesized proteins in the endoplasmic reticulum determine the transport pathways of soybean storage protein [22].

Surface Hydrophobicity

Surface hydrophobicity is an important factor that contributes to physicochemical functions such as solubility, emulsifying, and foaming abilities [23,24]. At 35% saturation of ammonium sulfate, A2B1a completely precipitated whereas the other proglycinins were completely soluble, indicating that A2B1a has the highest surface hydrophobicity among the five proglycinins. To compare the surface hydrophobicity of the other four proglycinins, they were subjected to hydrophobic columns, butyl and phenyl sepharoses. The higher the surface hydrophobicity, the stronger the surface interaction with the column resulting into slower elution. The order of elution of proglycinin was identical in both columns (Table 2). The results indicate that the order of hydrophobicity is as follows: A1bB2 > A3B4 ≥ A1aB1b > A5A4B3. The surface hydrophobicity was measured at pH 7.6. At pH 7.6, all proglycinins were soluble at $\mu = 0.5$, and the order of solubility at $\mu = 0.08$ was A5A4B3 > A3B4 > A1aB1b ≥ A1bB2 ≥ A2B1a. Therefore, there is no relationship between surface hydrophobicity and solubility of proglycinins.

Table 2. Elution time of proglycinin on hydrophobic column

| hydrophobic column | elution time (min) of proglycinin subunit | | | | |
|----------------------|---|-------|--------------------|------|--------|
| | A1aB1b | A1bB2 | A2B1a ^a | A3B4 | A5A4B3 |
| butyl sepharose 4FF | 48.7 | 54.3 | | 50.7 | 43.7 |
| phenyl sepharose 6FF | 67.0 | 74.4 | | 71.1 | 65.5 |

^aA2B1a was not subjected to the columns as it completely precipitated at 30% saturation of ammonium sulfate.

The surface hydrophobicities of various glycinin molecular species were assessed by the same hydrophobic columns, and the order of surface hydrophobicities was identical in both columns. The order was A3B4-glycinin > group II-glycinin > A5A4B3-glycinin > group I-glycinin (see Chapter I). This order was completely different from that of proglycinin, indicating that hexamer formation results in drastic change of surface hydrophobicities of glycinin subunits.

DSC Measurement

DSC has been used extensively to study thermodynamic properties of protein denaturation. The author performed an experiment on Microcal MC-2 ultrasensitive microcalorimeter (Figure 5).

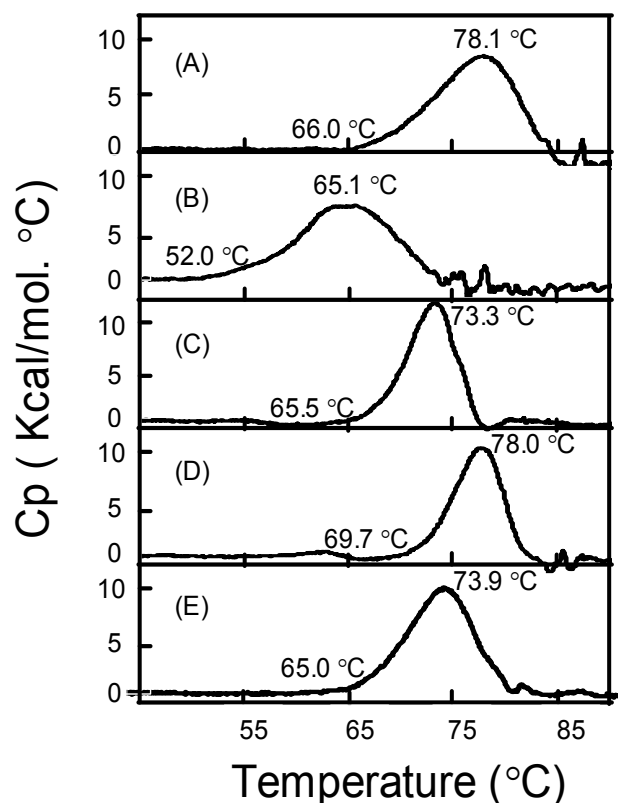


Figure 5. DSC scans of proglycinins. (A) A1aB1b, (B) A1bB2, (C) A2B1a, (D) A3B4, and (E) A5A4B3.

The thermal denaturation midpoint temperatures (T_m) and the denaturation starting temperature of proglycinins ranged from 65 to 78 °C and 52 to 70 °C, respectively. A1aB1b and A3B4 exhibited the highest T_m values while A1bB2 had the lowest. The order of T_m values was $A1aB1b \geq A3B4 > A5A4B3 \geq A2B1a > A1bB2$. It was noted that the shapes of exothermic peaks of A1aB1b and A1bB2 were broad, and those of A2B1a, A3B4, and A5A4B3 were sharp, indicating that the denaturation of A1aB1b and A1bB2 started at much lower temperatures than their T_m values. This resulted in the order of

denaturation starting temperature as follows: A3B4 > A1aB1b ≥ A2B1a ≥ A5A4B3 > A1bB2. The T_m values of A1aB1b and A3B4 were 78.1 and 78.0 °C, respectively, being very close to each other, but their denaturation starting temperatures were 66.5 and 69.7 °C, respectively. This indicates that A3B4 is the most stable proglycinin trimer among the five proglycinins. A2B1a and A5A4B3 exhibited very close T_m values (73.3 °C for A2B1a and 73.9 °C for A5A4B3) and denaturation starting temperatures (65.5 °C for A2B1a and 65.0 °C for A5A4B3). Both the T_m value and the denaturation starting temperature of A1bB2 were the lowest among the five proglycinins. Therefore, the order of the thermal stability of the five proglycinins is A3B4 > A1aB1b > A5A4B3 ≥ A2B1a > A1bB2. In the case of mature glycinins, the thermal stability of A3B4 glycinin was the lowest, 87.2 °C (see Chapter I). These show that the processing of proglycinin to mature form and the resultant hexamer formation result in stabilization of the molecules, but the degree of stabilization is variable among subunits.

Many factors were found to cause the differences in thermal stability. The author's group has pointed out five factors to be responsible for the difference in thermal stability between soybean β -conglycinin, β subunit (90.8 °C), and the core region of α' subunit (82.7 °C) [25]. Thus, (i) the total cavity volume is larger in α'_c ; (ii) the cluster of charged residues at the intermonomer interface is smaller in α'_c , and α'_c lacks intermonomer salt bridge; (iii) the solvent accessible surface is more hydrophobic in α'_c ; (iv) there are fewer proline residues in α'_c ; and (v) a loop region between helix 3 and strand J' in α'_c is longer. The combined contributions of these five factors should provide more hydrogen bonds in α'_c . To understand the structural reasons why individual proglycinins exhibit different thermal stability, X-ray crystallography of the proglycinins is required.

After heating to 110 °C and cooling to room temperature for the DSC experiment, precipitation of each proglycinin was observed. A3B4 and A1aB1b were found to be the least prone to precipitate, followed by A5A4B3, A2B1a, and A1bB2 subunits; A3B4 ≤ A1aB1b < A5A4B3 < A2B1a < A1bB2. Interestingly, A2B1a was very turbid, but no bulk precipitation was found. The degree of precipitation of proglycinins corresponds to the order of their thermal stability.

Emulsification

The emulsifying ability of proglycinins was analyzed by measuring particle size distributions after homogenization and sonication of samples with soybean oil at pH 7.6 and two ionic strengths ($\mu = 0.08$ and $\mu = 0.5$) (Figure 6). Proteins with a good emulsifying ability exhibit a small average particle size. Results obtained here indicate that all subunits exhibited a better emulsifying ability at $\mu = 0.5$ than at $\mu = 0.08$. The emulsifying ability of proglycinins can be ranked as follows: A1aB1b (3.9 and 10.5 μm) \geq A3B4 (4.0 and 11.7 μm) $>$ A5A4B3 (6.7 and 18.8 μm) $>$ A2B1a (9.9 and 19.4 μm) $>$ A1bB2 (14.9 and 47.3 μm). Values in parentheses indicate the average particle sizes at

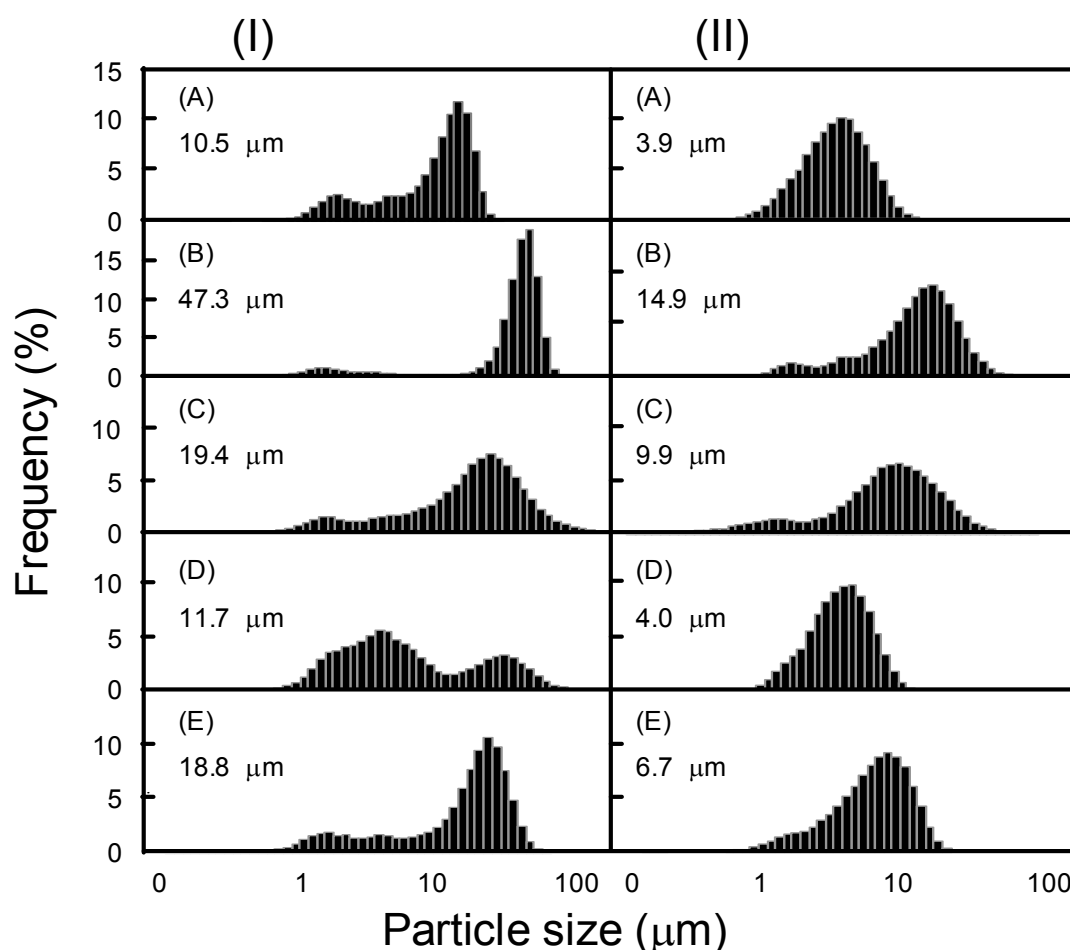


Figure 6. Particle size distributions of emulsion of proglycinins at ionic strengths of 0.08 (I) and 0.5 (II). (A) A1aB1b, (B) A1bB2, (C) A2B1a, (D) A3B4, and (E) A5A4B3.

high and low ionic strengths. Surface hydrophobicity was reported to be an important factor for physicochemical function such as the emulsifying ability, because it is related to the binding of a protein to oil [23,24]. However, the order of emulsifying abilities of proglycinins did not follow their order of surface hydrophobicities. A1bB2 had higher surface hydrophobicity among the five subunits but exhibited the poorest emulsifying ability. A2B1a had the highest surface hydrophobicity, but its emulsifying ability was only slightly better than that of A1bB2. In addition, suitable conformational change at the interface between oil and water, structural flexibility, is also one important factor [26]. Thermal stability is related to structural flexibility, but the order of the emulsifying abilities of proglycinins did not follow the order of their thermal stabilities.

Generally, a balance of hydrophobicity and hydrophilicity of a protein is a factor for its emulsifying ability [26]. Each glycinin subunit has 4–6 variable regions [17,27,28] that are rich in acidic residues. The longest one of the variable regions is located at the C-terminus of the acidic polypeptide and is called a hypervariable region that consists of 42, 29, 35, 70, and 103 residues corresponding to 8.8, 6.3, 7.5, 14.2, and 19.1% of the total amino acid residues [29] in A1aB1b, A1bB2, A2B1a, A3B4, and A5A4B3, respectively. Similarly, the N-terminus extension region of β -conglycinin α and α' consists of 125 and 141 residues corresponding to 23 and 24.4% of the total amino acid residues of α and α' , respectively. In previous studies, the author demonstrated that the hypervariable region of mature glycinin (see Chapter I) and the extension region of β -conglycinin [14] played a critical role in emulsifying ability. However, the order of contribution of variable region size as compared to the molecular size of proglycinin was A5A4B3 > A3B4 > A1aB1b > A2B1a > A1bB2, being not all consistent with the order of emulsifying ability of proglycinins. The effect on the emulsifying ability of the hypervariable region of proglycinins is different from that of mature glycinins and the extension region of β -conglycinin subunits. This might be due to the fact that the hypervariable region of proglycinin is located at the inside of the molecule, whereas the extension region and the hypervariable region are located at the N-terminus of β -conglycinin and C-terminus of acidic polypeptide of mature glycinin, respectively. Therefore, these regions can be more flexible than the hypervariable region in the proglycinins.

Thus, the emulsifying ability of proglycinin results from a combination of many factors such as surface hydrophobicity, solubility, flexible region (variable region), and structural stability that contribute to the balance of hydrophobicity and hydrophilicity and its suitable conformational change at the interface between oil and water.

Proglycinins have intrinsic solubility and have different physicochemical properties from each other. These characteristics are determined by many factors one of which might be the variable region. Recently, Utsumi's group succeeded in the X-ray crystallography of proglycinin A1aB1b [17] and mature glycinin A3B4 [28]. X-ray crystallography of the other subunits will further elucidate the molecular basis for these properties.

References

- [1] Utsumi, S., Kohno, M., Mori, T. and Kito, M. (1987). An alternate cDNA encoding glycinin A1aBx subunit. *J. Agric. Food Chem.* **35**, 210–214.
- [2] Utsumi, S., Kim, C. S., Kohno, M. and Kito, M. (1987). Polymorphism and expression of cDNAs encoding glycinin subunits. *Agric. Biol. Chem.* **51**, 3267–3273.
- [3] Choi, S.-K., Adachi, M. and Utsumi, S. (2002). Identification of the bile acid-binding region in the soy glycinin A1aB1b subunit. *Biosci., Biotechnol., Biochem.* **66**, 2395–2401.
- [4] Katsube, T., Kurisaka, N., Ogawa, M., Maruyama, N., Ohtsuka, R., Utsumi, S. and Takaiwa, F. (1999). Accumulation of soybean glycinin and its assembly with the glutelins in rice. *Plant Physiol.* **120**, 1063–1073.
- [5] Shirzadegan, M., Christie, P. and Seemann, J. R. (1991). An efficient method for isolation of RNA from tissue cultured plant cells. *Nucleic Acids Res.* **60**, 6055.
- [6] Cho, T.-J., Davies, C. S., Fisher, R. L., Turner, N. E., Goldberg, R. B. and Nielsen, N. C. (1989). Molecular characterization of an aberrant allele for the *Gy3* glycinin gene: A chromosomal rearrangement. *Plant Cell* **1**, 339–350.
- [7] Sanger, F., Nicklen, S. and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A* **74**, 5463–5467.
- [8] Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.
- [9] Kim, C. S., Kamiya, S., Kanamori, J., Utsumi, S. and Kito, M. (1990). High-level expression, purification and functional properties of soybean proglycinin from *Escherichia coli*. *Agric. Biol. Chem.*, **54**, 1543–1550.
- [10] Burnette, W. H. (1981). Western blotting: electrophoretic transfer of proteins from SDS-polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radiolabeled protein. *Anal. Biochem.* **112**, 95–203.
- [11] Green, A. A. and Hughes, W. L. (1955). *Methods in Enzymology*; Colowick, S. P., Kaplan, N. O., Eds; Academic Press: New York, Vol. 1, p 67.
- [12] Maruyama, N., Katsube, T., Wada, Y., Oh, M. H., Barba de la Rosa, A. P., Okuda, E., Nakagawa, S. and Utsumi, S. (1998). The roles of the N-linked glycans and extension regions of soybean β -conglycinin in folding, assembly and structural features. *Eur. J. Biochem.* **258**, 854–862.
- [13] Maruyama, N., Mohamed Salleh, M. R., Tkahashi, K., Yagasaki, K., Goto, H., Hontani, N., Nakagawa, S. and Utsumi, S. (2002). Structure-physicochemical function relationships of soybean β -conglycinin heterotrimers. *J. Agric. Food Chem.* **50**, 4323–4326.
- [14] Maruyama, N., Sato, R., Wada, Y., Matsumura, Y., Goto, H., Okuda, E., Nakagawa, S. and Utsumi, S. (1999). Structure-physicochemical function relationships of soybean β -conglycinin constituent subunits. *J. Agric. Food Chem.* **47**, 5279–5284.
- [15] Nielsen, N. C., Dickinson, C. D., Cho, T. J., Thanh, V. H., Scallan, B. J., Fischer, R. L., Sims, T. L., Drews, G. N. and Goldberg, R. B. (1989). Characterization of the glycinin gene family in soybean. *Plant Cell* **1**, 313–328.
- [16] Utsumi, S., Gidamis, A. B., Mikami, B. and Kito, M. (1993). Crystallization and preliminary X-ray crystallographic analysis of the soybean proglycinin expressed in *Escherichia coli*. *J. Mol. Biol.* **233**, 177–178.
- [17] Adachi, M., Takenaka, Y., Gidamis, A. B., Mikami, B. and Utsumi, S. (2001). Crystal structure of soybean proglycinin A1aB1b homotrimer. *J. Mol. Biol.* **305**, 291–305.

- [18] Kim, C. S., Kamiya, S., Sato, T., Utsumi, S. and Kito, M. (1990). Improvement of nutritional value and functional properties of soybean glycinin by protein engineering. *Protein Eng.* **3**, 725–731.
- [19] Katsube, T., Gidamis, A. B., Kanamori, J., Kang, I. J., Utsumi, S. and Kito, M. (1994). Modification tolerability of the hypervariable region of soybean proglycinin. *J. Agric. Food Chem.*, **42**, 2639–2645.
- [20] Maruyama, N., Fukuda, T., Saka, S., Inui, N., Kotoh, J., Miyagawa, M., Hayashi, M., Sawada, M., Moriyama, T. and Utsumi, S. (2003). Molecular and structural analysis of electrophoretic variants of soybean seed storage proteins. *Phytochemistry* **64**, 701–708.
- [21] Utsumi, S. (1992). Plant food protein engineering. In *Advances in Food and Nutrition Research*; Kinsella, J. E., Ed.; Academic Press: SanDiego, CA, pp 89–208.
- [22] Mori, T., Maruyama, N., Nishizawa, K. Higasa, T., Yagasaki, K., Ishimoto, M. and Utsumi, S. (2004). The composition of newly synthesized proteins in the endoplasmic reticulum determines the transport pathways of soybean seed storage proteins. *Plant J.* **40**, 238–349.
- [23] Nakai, S. (1983). Structure-function relationships of food proteins with an emphasis on the importance of protein hydrophobicity. *J. Agric. Food Chem.* **31**, 676–683.
- [24] Nakai, S. and Li-Chan, E. (1988). *Hydrophobic Interaction in Food Systems*; CRC Press: Boca Raton, FL.
- [25] Maruyama, Y., Maruyama, N., Mikami, B. and Utsumi, S. (2004). Structure of the core region of the soybean β -conglycinin α' subunit. *Acta Crystallogr.* **D60**, 289–297.
- [26] Damodaran, S. (1997). Protein-stabilized forms and emulsions. In *Food Proteins and Their Applications*; Damodaran, S., Paraf, A., Eds.; Marcel Dekker: New York, pp 57–110.
- [27] Lawrence, M. C., Izard, T., Beuchat, M., Blagrove, R. J. and Colman, P. M. (1994). Structure of phaseolin at 2.2 Å resolution implications for a common vicilin/legumin structure and the genetic engineering of seed storage proteins. *J. Mol. Biol.* **238**, 748–778.
- [28] Adachi, M., Kanamori, J., Masuda T., Yagasaki K., Kitamura, K., Mikami, B. and Utsumi, S. (2003). Crystal structure of soybean 11S globulin: Glycinin A3B4 homohexamer. *Proc. Natl. Acad. Sci. U.S.A* **100**, 7395–7400.
- [29] Utsumi, S., Matsumura, Y and Mori, T. (1997). Structure-function relationships of soy proteins. In *Food Proteins and Their Applications*; Damodaran, S., Paraf, A., Eds.; Marcel Dekker: New York, pp 257–291.

Chapter III

Protein Engineering of Soybean Proglycinin C-Terminus for Improving Physicochemical Properties

Emulsifying ability of a protein is one of its significant functional properties in relation to its application in food systems and it depends basically on two effects: (1) a substantial decrease in the interfacial tension due to the adsorption of the protein at the oil-water interface and (2) the electrostatic, structural and mechanical energy barrier caused by the interfacial layer that opposes destabilization processes [1,2]. The relationship between physicochemical characteristics, interfacial behaviors, and emulsifying abilities of proteins was extensively investigated. It was suggested that surface hydrophobicity is one of the characteristics of the protein most likely to define its surface behaviors and consequently its emulsifying properties. Though such hypotheses have been suggested for the characterization of the emulsifying properties of proteins, improvement of emulsifying ability of a protein has been limited. Due to the limitation for the improvement of emulsifying ability, but a high demand for utilization of soybean protein in food products, it is important to improve the emulsifying property of soybean proteins by protein engineering.

The emulsifying activity of soy proteins has already been extensively studied [3–7]. Native glycinin was reported to exhibit surface behavior and, consequently, foaming and emulsifying properties, which are limited by its closely packed globular conformation, low surface hydrophobicity, and low molecular flexibility [8,9]. Its solubility and surface-active properties were improved when the oligomeric structure of glycinin was appropriately dissociated with simultaneous unfolding of the acidic and basic polypeptide chains [10]. It was later confirmed that the acidic polypeptide of soybean glycinin has good emulsifying ability [5]. On the other hand, analysis of the individual β -conglycinin subunits showed that the extension regions of α and α' subunits that are rich in negatively charged residues confere better solubility and emulsifying ability while the core regions

determine the thermal stability [4,11]. This is consistent with the observation that the negatively charged variable region IV of glycinin also contributed to its solubility and emulsifying ability, although the extent is lower than that of the extension region (Chapter I). The presence or absence of the highly negatively charged extension region in β -conglycinin and the length of the highly negatively charged variable region IV in glycinin should be therefore important determinants of solubility and emulsifying property. However, previously it was found that addition of variable region IV (42 aa) of proglycinin A1aB1b to its variable regions II, III and C-terminus, and α' extension region (141 aa) of β -conglycinin α' subunit to proglycinin A1aB1b variable region IV did not improve emulsifying property of the protein. Differently, addition of α' extension region to A1aB1b C-terminus improved emulsion property and stability remarkably [12]. Change the position of α' extension region from N- to C-terminus of β -conglycinin α' subunit showed improved emulsion stability. Tandang et al. [12] suggested that C-terminus of proglycinin and α' extension polypeptide are good position and polypeptide, respectively, for the improvement of emulsifying property of the proglycinin. α' Extension region is long and rich in charged residues, especially negatively charged residues. Then, what structural characteristics: length, hydrophilicity, or charged residues in the region is important for the improvement? It is noted that α extension (125 aa) and A5A4B3 hypervariable (103 aa) regions are also long and rich in charged residues. However, hydrophobicity profiles of α' and α extension regions and A5A4B3 hypervariable region are different.

At present study, to investigate the structural characteristics, the author introduced α' and α extension regions to A2B1a C-terminus, and α extension and A5A4B3 hypervariable (A4IV) regions to A1aB1b C-terminus. To evaluate the influence of positively or negatively charged residues at the C-terminus, she also introduced an oligopeptide composed of twenty negatively or positively charged residues to the A1aB1b C-terminus. All the mutants were expressed in *Escherichia coli* and physicochemical properties such as thermal stability, surface hydrophobicity, solubility and emulsifying property of the mutants were studied.

Materials and Methods

Construction of Expression Plasmids for Mutants

Schematic representations of proglycinin wild types (WT) and their mutants are shown in Figure 1. To construct expression plasmids for the mutants, the expression plasmids pEA1aB1b [13], pEA2B1a, pEA5A4B3 (see Chapter II), pEC α , pEC α' [11] were used as templates for PCR. The different primers used in amplifying the desired mutant cDNAs by PCR using KOD plus (TOYOBO) are as follows: pEA1aB1b as template for construction pEA1aB1b α' , pEA1aB1b α , pEA1aB1bA4IV (cDNA encoding

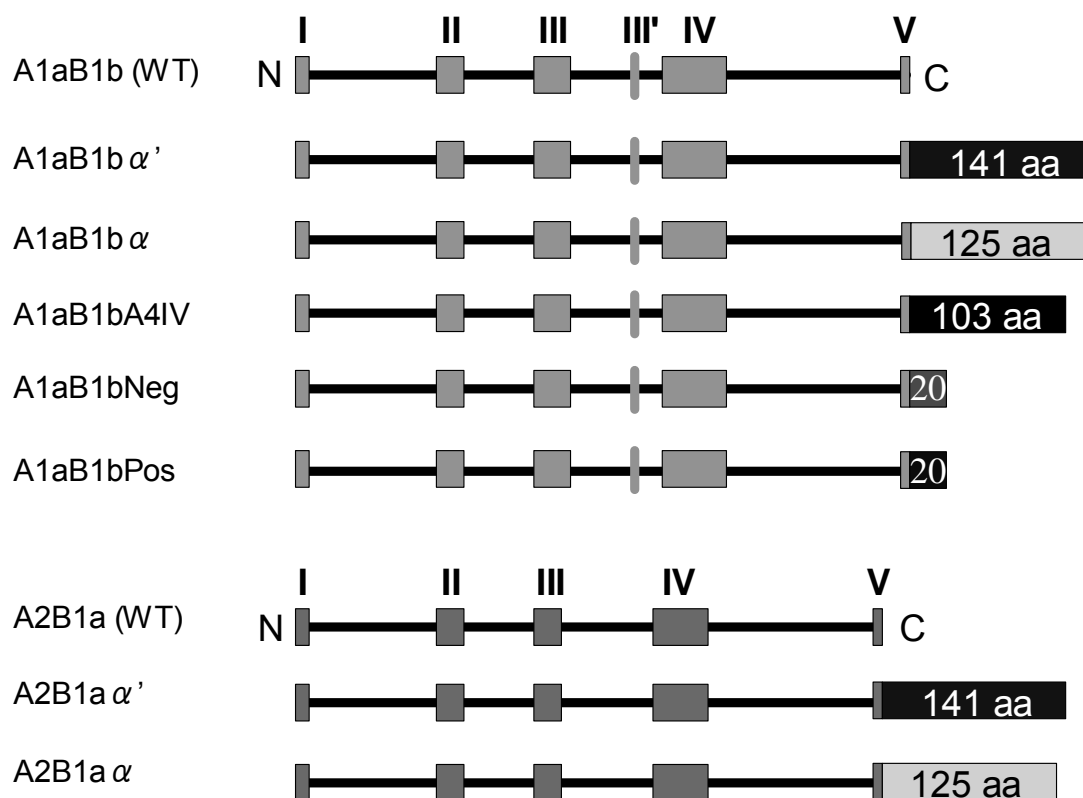


Figure 1. Schematic diagram of proglycinin WT and their mutants. Roman numeral numbers name the disordered and variable regions of A1aB1b and A2B1a, respectively, shown in gray. Arabic numeral numbers indicate the amount of amino acid extended to the C-termini (Pos, Neg, A5A4B3 hypervariable, α and α' extension regions consisted of 20 positively, 20 negatively, 103, 125 and 141 amino acid residues, respectively).

A1aB1b–pET-21d), 5'-TAG AAT TCC GGA TCC GAA TTC GAG CTC-3' and 5'- AGC CAC AGC TCT CTT CTG AGA CTC C-3'; pEα' as template for construction pEA1aB1bα' and pEA2B1αα' (cDNA encoding α' extension region), 5'-GTG GAG GAA GAA GAA GAA TGC GAA GAA GGT C-3' and 5'-TGG TTC TCT TTG AGA CTC AGA ACC TTC-3'; pEα as template for construction pEA1aB1bα and pEA2B1αα (cDNA encoding α extension region), 5'-GTG GAG AAA GAA GAA TGT G-3' and 5'-TAA CTC AGA ATC TTC ACT TTC TTC GCT-3'; pEA5A4B3 as template for construction of pEA1aB1bA4IV (cDNA encoding A5A4B3 hypervariable region), 5'-AAG TGG CAA GAA CAA CAA GAT GAA G-3' and 5'-TCT TGT CTC GCA TCC TCT TTC ACG T-3'; pEA2B1a as template for construction pEA2B1αα and pEA2B1αα' (cDNA encoding A2B1a–pET-21d), 5'-TAG AAG CTT GCG GCC GCA CTC GAG CAC-3' and 5'-AGC CAC AGC TCT CCT CTG AGA CTC-3'; pEA1aB1b as template for pEA1aB1bNeg, 5'-**GAA GAA GAA GAA GAA GAG GAG GAG GAG GAG TAG** AGC CCT TTT TGT ATG GAT CCG-3' and 5'- **GTC GTC GTC ATC ATC ATC ATC ATC ATC ATC** AGC CAC AGC TCT CTT CTG AGA CTC-3'; pEA1aB1b as template for pEA1aB1bPos, 5'-**CGC CGT AAG CGT AAA CGT CGT CGT CGT AAA TAG** AGC CCT TTT TGT ATC GAT CCG-3' and 5'-**TTT CTT CCG CTT GCG CTT CTT CTT CTT GCG** AGC CAC AGC TCT CTT CTG AGA CTC-3' (italic and bold letters are stop codon and condons for the insertion of charged amino acid residues, respectively).

For amplifying cDNAs encoding for A1aB1b–pET-21d and A2B1a–pET-21d, 30 cycles of denaturation at 94 °C for 15 s, annealing at 60 °C for 30 s, and elongation at 68 °C for 10 min was used. The cDNAs encoding for A5A4B3 hypervariable region, α and α' extension regions were amplified by 25 cycles of 94 °C for 15 s, annealing at 60 °C for 30 s, and elongation at 68 °C for 3 min. The resulting fragments corresponding to A5A4B3 hypervariable, α and α' extension regions were phosphorylated and blunted before ligations with the corresponding vectors to construct the expression plasmid pEA1aB1bα', pEA1aB1bα, pEA1aB1bA4IV, pEA2B1αα', pEA2B1αα. The reaction conditions for cDNAs of A1aB1bNeg–pET-21d and A1aB1bPos–pET-21d were 30 cycles of denaturation at 94 °C for 15 s, annealing at 65 °C for 40 s, and elongation at 72 °C for 7

min, and extended for 10 min at 72 °C. The resulting fragments corresponding to A1aB1bNeg–pET-21d and A1aB1bPos–pET-21d were phosphorylated and ligated to produced pEA1aB1bNeg and pEA1aB1bPos.

Protein Expression

The expression plasmids were transformed into *E. coli* expression hosts HMS174(DE3), BL21(DE3), AD494(DE3), Origami(DE3). Culture and expression conditions of WT A1aB1b and A2B1a was as described previously in Chapter II. The cells were grown at 37 °C. Expression was induced when A_{600} reached 0.4 to 0.6 with 1 mM isopropyl-1-thio- β -D-galactoside (IPTG), at 18 and 20 °C. After cultivation, cells were harvested by centrifugation at 6000 rpm for 15 min at 4 °C and stored until used. Aliquots of the cells were analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE) [14] using 11% gel as described previously [15]. Expressed recombinant proteins were identified based on their expected sizes and confirmed by western blotting [16] using anti-glycinin antibody followed by goat-rabbit IgG-alkaline phosphatase conjugate (Promega).

Purification of Mutants

All purification steps were carried out at 4 °C and centrifugation were at 6000 rpm for 20 min unless otherwise stated. The basic buffer for all purification steps was buffer A (35 mM potassium phosphate, pH 7.6, 1 mM EDTA, 10 mM 2-mercaptoethanol, 0.1 mM (*p*-amidinophenyl)-methylsulfonyl fluoride, 1 μ g/ml pepstatin A, 1 μ g/ml leupeptin). Ammonium sulfate fractionation followed the procedure of Green and Hughes [17].

A1aB1b and A2B1a wild types were purified as described previously (see Chapter II). Frozen cells containing A1aB1b α' and A1aB1b α were resuspended in buffer B (buffer A containing 0.4 M NaCl); A1aB1bA4IV, A1aB1bNeg and A1aB1bPos were resuspended in buffer C (buffer A containing 1.0 M NaCl); A2B1a α' and A2B1a α were resuspended in buffer D (identical to buffer C but containing 35 mM sodium sulfate instead of potassium sulfate) (Table 1) at a density of 40 g/l buffer and lysed by sonication on an ice

bath. Insoluble matter was removed by centrifugation. Expressed mutants were fractionated using ammonium sulfate: 30% for A1aB1b α' , A1aB1b α , A2B1a α' , A2B1a α , and A1aB1bNeg; 35% for A1aB1bA4IV and A1aB1bPos. The precipitate was removed by centrifugation, and the soluble fraction containing recombinant proteins was applied on a Toyopearl (Butyl-650M) (TOSOH) column (2.6 cm x 20 cm) equilibrated with buffer B containing 30% ammonium sulfate. Elution was carried out with a linear gradient (800 ml) from 30 to 0% ammonium sulfate in buffer B for A1aB1b α' and A1aB1b α , and in buffer C for the other mutants. Fractions containing A1aB1b α' and A1aB1b α were dialyzed against buffer E (buffer A containing 0.15 M NaCl) and clarified by centrifugation. They were applied onto Mono Q HR 10/10 column (Pharmacia Biotech) equilibrated with buffer C. Elution was performed with a linear gradient from 0.15 to 0.6 M NaCl in buffer A over a period of 120 min at 2 ml/min. Fractions containing A1aB1bA4IV, A1aB1bNeg, A2B1a α' and A2B1a α were concentrated by VIVASPIN 20 (VIVASCIEN, Japan) and applied on a gel filtration column (Hi-Prep 26/60 Sephacryl S-300 HR) using buffer C as a mobile phase. However, fractions containing A1aB1bPos were concentrated by dialysis in buffer C containing 75% ammonium sulfate. The precipitate containing A1aB1bPos was dissolved in 2 ml buffer C and then it was applied on the same gel filtration column.

Protein Measurement

The amount of protein in the samples was determined using a Protein Assay Rapid Kit (Wako) with bovine serum albumin as a standard.

Analysis of Self-Assembly Into Trimers

Self-assembly of individual mutants was analyzed using Hi-Prep 16/60 Sephacryl S-300 HR column (Pharmacia Biotech) as described previously in Chapter II. All the samples used were 500 μ l of 0.25 mg/ml in buffer F [35 mM sodium phosphate, pH 7.6, 0.4 M NaCl, 1 mM EDTA, 0.1 mM (*p*-amidinophenyl)-methylsulfonyl fluoride, 1 μ g/ml pepstatin A, 1 μ g/ml leupeptin, 0.02% NaN₃, 10 mM 2-mercaptoethanol] except

A1aB1bPos that was in buffer G [identical to buffer F but containing 1.0 M NaCl]. The mobile phase using for all the samples was buffer F and flow rate was 0.5 ml/min.

Differential Scanning Calorimetry (DSC) Measurement

DSC measurement of samples was carried out as described previously in Chapter II using 1 mg/ml of the sample in buffer F. Scanning was recorded using Microcal MC-2 Ultra Sensitive Microcalorimeter (Micro Cal Inc., Northampton, MA) at a rate of 1 °C/min.

Surface Hydrophobicity

Surface hydrophobicities of samples were analyzed as described previously in Chapter II using butyl and phenyl sepharose (Amersham Bioscience, Sweden), and 500 μ l, 0.25 mg/ml of the samples in buffer K (buffer F containing 35% saturation of ammonium sulfate). The proteins were eluted with a linear gradient of ammonium sulfate 35 to 0% in buffer F over a period of 55 min and extended running with buffer F for 45 min at a flow rate of 0.25 ml/min.

Solubility Analysis as a Function of pH

Except A1aB1bPos, all samples were dialyzed against buffer H [10 mM sodium phosphate, pH 7.6, 0.5 M NaCl, 1 mM EDTA, 0.1 mM (*p*-amidinophenyl)-methylsulfonyl fluoride, 1 μ g/ml pepstatin A, 1 μ g/ml leupeptin, 0.02% NaN₃, 10 mM 2-mercaptoethanol]. Experimental conditions and analysis were done as described previously in Chapter II. A1aB1bPos that precipitated in buffer H was avoided by dialyzed against buffer I [5 mM sodium phosphate, pH 7.6, μ = 0.014]. For the analysis at the same ionic strength and pH with the other samples, appropriate amount of 10 and 50 mM of citrate buffer (pH 2.5 to 6.5), sodium phosphate buffer (pH 5.5 to 7.8) and ammonium buffer (pH 7.2 to 9.0) was added. The final ionic strength 0.08 and 0.5 of the sample was adjusted by adding appropriate amount of buffer J (buffer I containing 3.422 M NaCl). The sample (0.8 mg/ml) was then analyzed same procedure as the other samples.

Emulsifying Property

The emulsifying properties of samples were analyzed as described previously [12] using 0.5 mg/ml of proteins in buffer F and buffer L (identical to buffer H but containing 0.05 M NaCl) for $\mu = 0.5$ and 0.08, respectively.

Results and Discussion

Expression, Purification and Self-Assembly of Individual Mutants

AD494(DE3) was a good expression host for WT A1aB1b and A2B1a. The expressions of these proglycinins as soluble proteins were estimated at about 15% of the total protein as shown in Chapter II. However, the good expression host of all mutants was different according to the individual mutant. After assessment of the expression efficiency and solubility of the recombinant proteins, the most suitable *E. coli* strain for each plasmid was selected to be as follow: HMS174(DE3) for A1aB1b α' ; BL21(DE3) for A1aB1bA4IV and A1aB1bPos; AD494(DE3) for A1aB1b α , A2B1a α' and A2B1a α ; and, Origami(DE3) for A1aB1bNeg. To achieve a high level expression of a soluble mutant, LB medium containing NaCl from 0.4 to 0.6 M was used and the most suitable *E. coli* strain harboring each individual expression plasmid was cultured at 20 °C for 40 h (Table 1). However, A1aB1bPos was cultured at 18 °C and harvested after overnight culture. In optimizing expression conditions, the author found that it was easier to get soluble expression of mutant A1aB1bs than mutant A2B1as being consistent with the expression property of their WTs as shown in Chapter II.

The solubility of recombinant proteins from individual expression plasmids in *E. coli* cells was analyzed by SDS–PAGE after sonication (Figure 2). On the basis of band intensity, the solubility of all the mutants was estimated to be > 80%. These were confirmed by western blotting (data not shown) using purified antiglycinin antibody.

Table 1. Host strain and culture condition for mutants

| Mutants | Host cell | Culture conditions for protein expression | | | | Sonication buffer | |
|------------------|--------------|---|-------------|------------------|--------------|------------------------|---------------------|
| | | Medium | Salt (NaCl) | Temperature (°C) | Duration (h) | Buffer (pH 7.6) | Salt content (NaCl) |
| A1aB1b α' | HMS174(DE3) | LB | 0.4M | 20 | 40 | 35mM KPi ^a | 0.4M |
| A1aB1b α | AD494(DE3) | LB | 0.4M | 20 | 40 | 35mM KPi | 0.4M |
| A1aB1bA5IV | BL21(DE3) | LB | 0.4M | 20 | 40 | 35mM KPi | 1.0M |
| A1aB1bNeg | Origami(DE3) | LB | 0.5M | 20 | 40 | 35mM KPi | 1.0M |
| A1aB1bPos | BL21(DE3) | LB | 0.5M | 18 | 20 | 35mM KPi | 1.0M |
| A2B1a α' | AD494(DE3) | LB | 0.6M | 20 | 40 | 35mM NaPi ^b | 1.0M |
| A2B1a α | AD494(DE3) | LB | 0.4M | 20 | 40 | 35mM NaPi | 1.0M |

^aKPi-potasium phosphate buffer.

^bNaPi-sodium phosphate buffer.

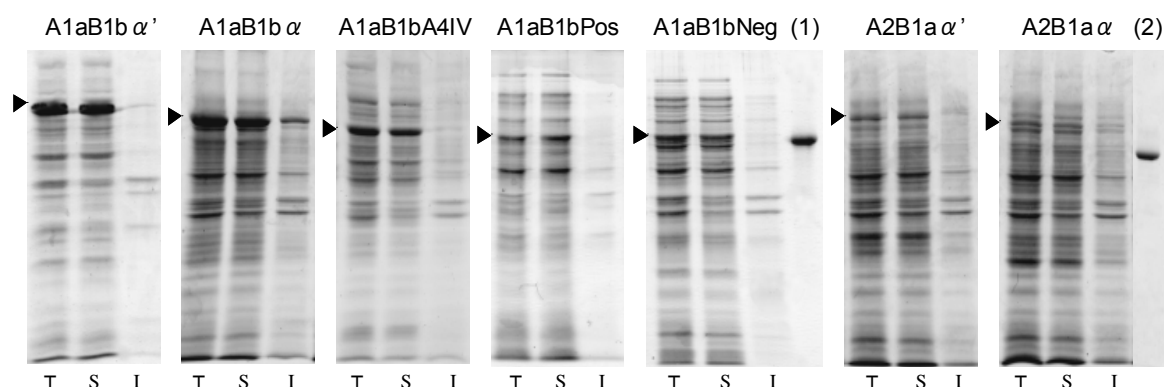


Figure 2. SDS-PAGE (11% acrylamide) profile of the protein expression of proglycinin mutants. The arrow indicates the position of the expressed mutants. T, total protein after cell extraction; S, soluble fraction; I, insoluble fraction. (1), purified wild type A1aB1b; (2), purified wild type A2B1a.

After purification, the purity of the proteins was assessed by SDS-PAGE (Figure 3) and found to be > 95%. The author found that the mobility of all mutants corresponded to their expected molecular sizes. Since all mutants were resulted from an increase in molecular size caused by addition of α' extension, α extension, A5A4B3 hypervariable region, negatively/positively charged oligopeptide corresponding to 141, 125, 103 and 20

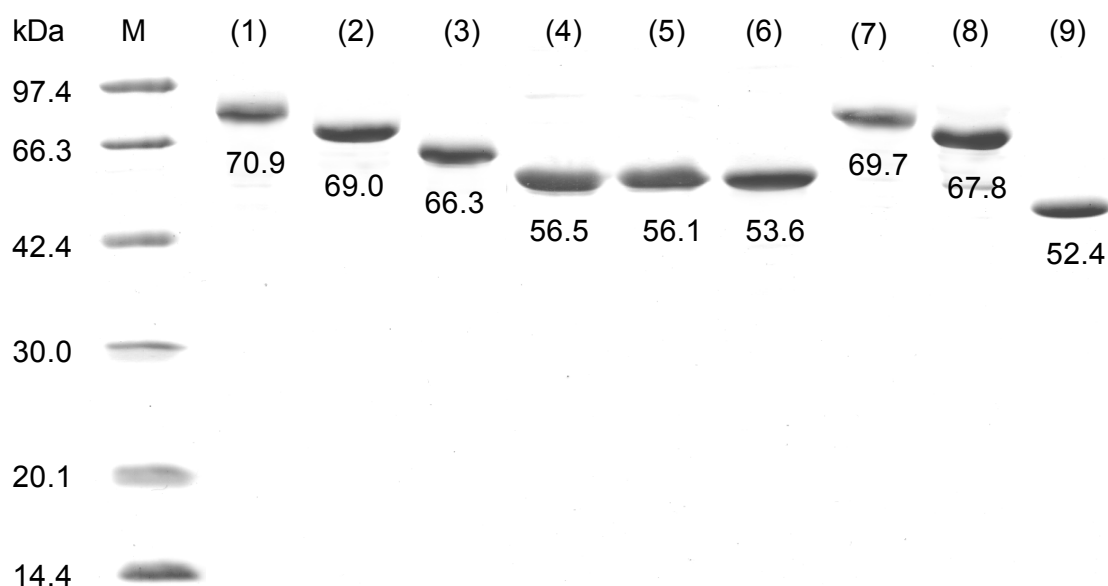


Figure 3. SDS-PAGE (11% acrylamide) analysis of the purified proglycinins and their mutants. (1), A1aB1b α' ; (2), A1aB1b α ; (3), A1aB1bA4IV; (4), A1aB1bPos; (5), A1aB1bNeg; (6), A1aB1b wild type; (7), A2B1a α' ; (8) A2B1a α , (9), A2B1a wild type; M, marker. The Arabic numeral indicate the molecular mass in kDa.

amino acid residues, respectively, to C-terminus region, it was necessary to test if the mutants form proper quaternary structures as the WT. To investigate, individual purified mutant and WT samples were subjected to gel filtration chromatography using Hi-Prep 16/60 Sephacryl S-300 HR column at pH 7.6 and $\mu = 0.5$. The elution times were 96.1, 96.6, 97.6, 98.4, 103.6, 112.6, 122.7, 123.1 and 130.1 for A1aB1b α' (70.9 kDa), A1aB1b α (69.0 kDa), A2B1a α' (69.7 kDa), A2B1a α (67.8 kDa), A1aB1bA4IV (66.3 kDa), A1aB1bNeg (56.1 kDa), A1aB1bPos (56.5 kDa), A1aB1b (53.6 kDa) and A2B1a (52.4 kDa), respectively (Figure 4). Proglycinins WT A1aB1b and A2B1a were found to be trimers as shown in Chapter II. All mutants were eluted from gel filtration column according to their molecular masses. A2B1a α' was seven amino acid residues bigger than A1aB1b α , it eluted slower. However, this phenomenon was not surprised since WT A2B1a had the most compact structure and eluted slower than the other proglycinin subunits as shown in Chapter II. Therefore, folding property of mutant A2B1as should be the same with that of WT A2B1a and should adapt theses elution behaviors. A1aB1bPos

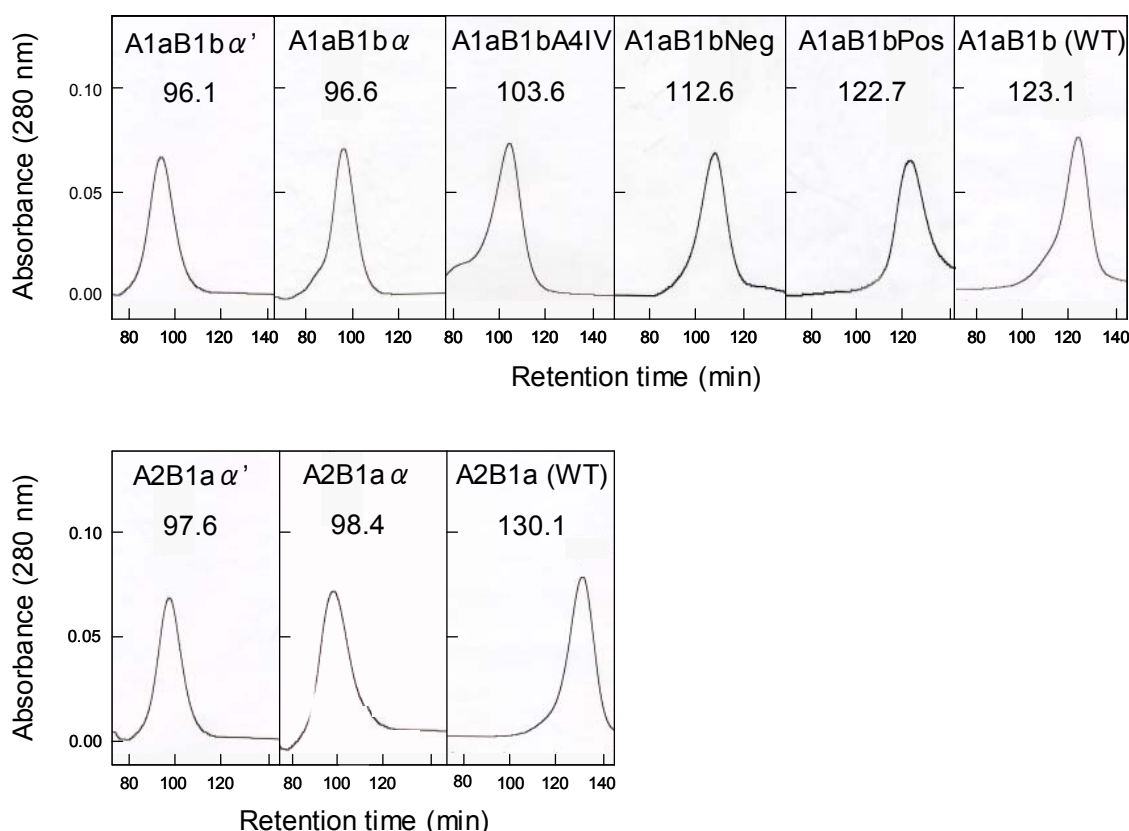


Figure 4. Gel filtration profile of proglycinins and their mutants. The purified proglycinin WT and mutants were subjected to chromatography using a sephacryl S-300 HR column.

was 0.4 kDa bigger than that of A1aB1bNeg, thus it should elute at least at the same time with that of A1aB1bNeg, but differently it was eluted at 10 min later and was only 0.4 min before WT A1aB1b. This phenomenon maybe due to that high positively charged oligopeptide at the C-terminus of A1aB1bPos interacted with variable regions of the mutant that was high negatively charged residues, and especially the interaction of a C-terminus of a monomer with variable regions II and IV of another monomer since after trimers formation the C-terminus and variable regions II and IV were nearby each other after trimer formation as show in Figure 5. This interaction brought about aggregation at moderate salt concentration buffer. In fact, these phenomenon were clearly observed during purification of this protein. This protein could be purified only when high salt buffer containing 1.0 M NaCl was used (data not shown). In case A1aB1bNeg the

negatively charged oligopeptide at the C-terminus might repulse from the negatively charged of the variable regions brought about the elution of this protein similar to a 58.3 kDa-mutant (IV-4+V-4) (see Chapter IV). As a result, A1aB1bNeg eluted a bit faster and A1aB1bPos eluted a bit slower than their normal sizes in gel filtration that made the difference of the two proteins with similar molecular sizes eluted at 10 min difference. Therefore, all the mutants were assessed to be trimers as their WTs.

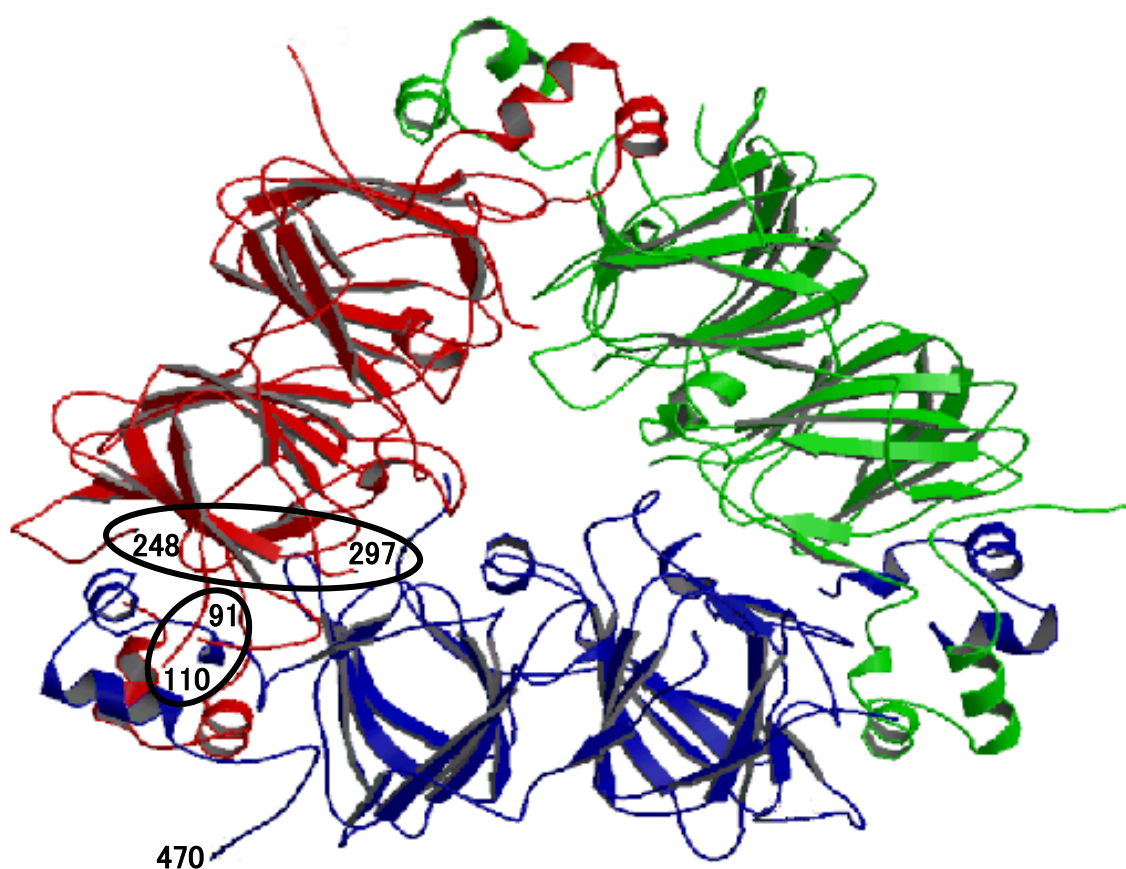


Figure 5. A ribbon diagram of proglycinin A1aB1b homotrimers structure (Adachi et al. [18]; 1FXZ) showing each monomer in red, blue and green. The Arabic numerals in each circle indicated a residue number before the start and after the end of disordered region II (residues 92-109) and IV (residues 249-296) of a red monomer. The Arabic numeral 470 indicated a residue before starting the disordered region V (C-terminus) of a blue monomer.

DSC Measurement

DSC analysis was studied to compare thermal stability of mutants and their WT (Figure 6). Comparing T_m values of mutant A1aB1bs and mutant A2B1as with that of the WT, indicated that A1aB1b α' , A1aB1b α and A1aB1bNeg exhibited a bit higher T_m values than that of the WT A1aB1b whereas A1aB1bA4IV and A1aB1bPos, and A2B1a α' and A2B1a α exhibited a bit lower T_m values than that of the WT A1aB1b and WT A2B1a, respectively. These results were consistent with a previous study that some mutant A1aB1bs gave higher T_m values and some gave lower T_m values than that of the WT [12].

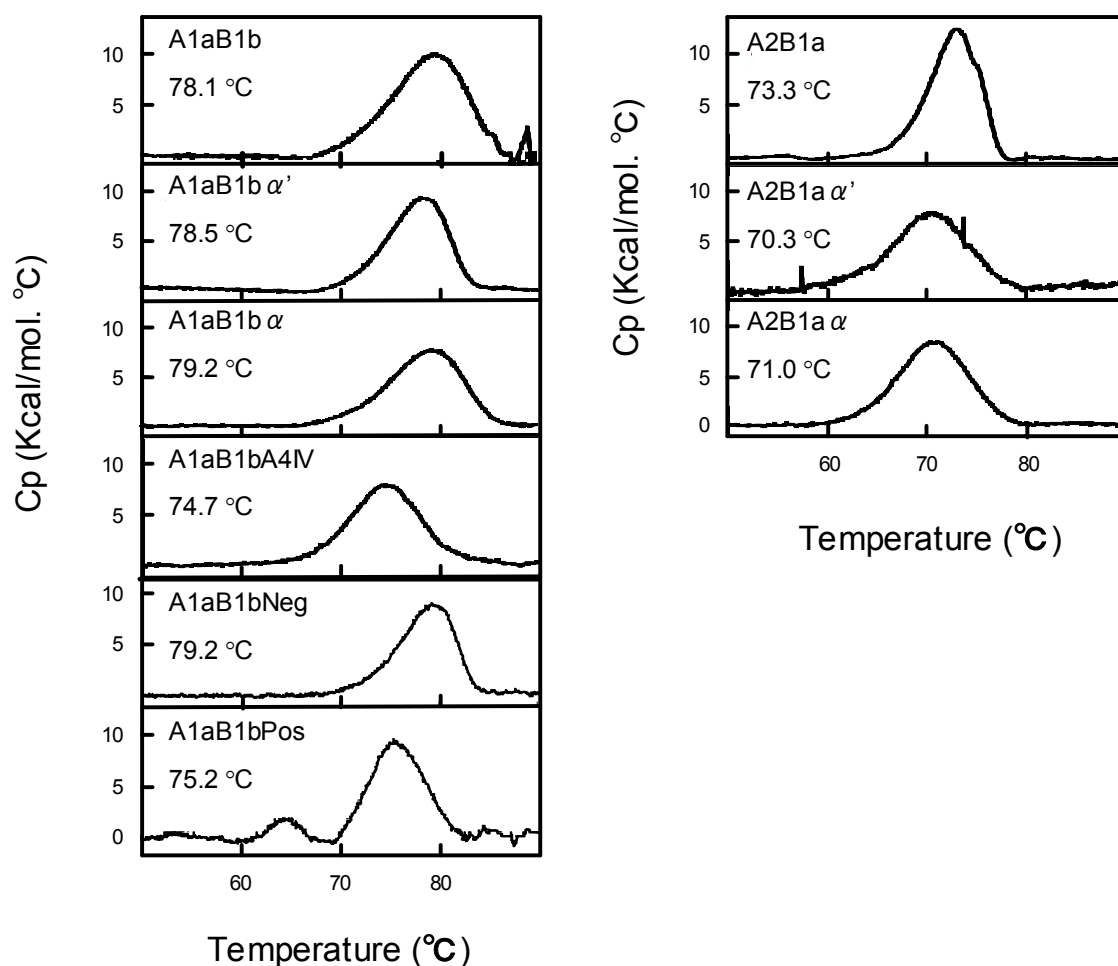


Figure 6. DSC scans of proglycinins and their mutants.

These results demonstrate that the addition of the α' and α extension, A5A4B3 hypervariable regions and negatively/positively charged twenty amino acid residues to the C-termini of proglycinins does not greatly affect the thermal denaturation points of these proglycinins. Modification in this study seems to be acceptable despite the drastic insertions. These results indicate that the tertiary structure of all the modified proteins constructed here can be assumed to fold like their WT_s.

Surface Hydrophobicity

Surface hydrophobicity of proteins is related to some of their physicochemical properties such as emulsifying and solubility [19,20]. Kato and Nakai [21] described that *cis*-parinaric acid (CPA) is more suitable for the measurement of surface hydrophobicity than 1-anilino-8-naphthalenesulfonate (ANS). Hayakawa and Nakai [22] reported that hydrophobicity measured by phenyl sepharose column chromatography and ANS correlated well to the protein insolubility, whereas no significant correlation was observed between that by CPA and insolubility. They suggested that the aromatic hydrophobicity may play a more important role in protein solubility than the aliphatic hydrophobicity. Therefore, the author employed two columns of phenyl and butyl sepharose. With this analysis, the longer the elution time is, the higher the surface hydrophobicity will be of the sample. Results (Table 2) showed that surface hydrophobicities of all the mutants accessed by the two columns were lower than that of their wild types except that of A1aB1bPos that was higher and eluted at 7 and 5.5 min later from butyl and phenyl sepharose, respectively.

Table 2. Elution time of proglycinin and mutant on hydrophobic column

| hydrophobic column | elution time (min) | | | | | | | |
|--------------------|--------------------|------------------|-----------------|------------|-----------|-----------|-----------------|----------------------------------|
| | A1aB1b (WT) | A1aB1b α' | A1aB1b α | A1aB1bA4IV | A1aB1bNeg | A1aB1bPos | A2B1a α' | A2B1a α (WT) ^a |
| butyl sepharose | 48.7 | 45.5 | 45.6 | 45.6 | 43.5 | 55.7 | 57.3 | 57.7 |
| phenyl sepharose | 67.0 | 64.7 | 64.8 | 64.9 | 62.4 | 72.5 | 72.8 | 73.4 |

^a A2B1a was not subjected to the columns as it completely precipitated at 30% saturation of ammonium sulfate.

Surface hydrophobicity of A1aB1bNeg is the poorest among all and followed by A1aB1b α' = A1aB1b α = A1aB1bA4IV < A1aB1b (WT) < A1aB1bPos < A2B1a α' \leq A2B1a α < A2B1a. Decrease surface hydrophobicity of the mutants might be due to the introduction of highly hydrophilic polypeptides to the C-terminus, which are 62.4, 60.8, 63.1 and 100% hydrophilicity for α' extension region, α extension region, A5A4B3 hypervariable region and twenty negatively charged residues, respectively. However, introduction of a fully positively charged oligopeptide did not reduce the surface hydrophobicity but increased instead. This might be due to the reason that in form of trimers, C-terminus of a monomer located near by disordered region II and IV of the other monomer (Figure 5). Both disordered regions II and IV were at an IE face [18]. The two disordered regions, especially disordered region IV, are highly in negatively charged residues. Therefore, there might be interaction between the negative charge of disordered region IV and the positive charged oligopeptide that was introduced to the C-terminus. The interaction might lead the IE face which is high in hydrophobic surface [18] to be higher especially in the presence of high ammonium sulfate concentration. Hence, resulted in increase the total surface hydrophobicity of the mutant A1aB1bPos.

Solubility as a Function of pH

Solubility is a fundamental physicochemical property of food proteins. The author measured the solubility of all mutants at high ($\mu = 0.5$) and low ($\mu = 0.08$) ionic strengths (Figure 7). Results (Figure 7A) showed that at $\mu = 0.5$, the solubility of A1aB1b α' and A1aB1bA4IV were similar to that of their WT. However, the solubility profiles of A1aB1b α' and A1aB1bA4IV shifted to lower pHs at $\mu = 0.08$. At this ionic strength, A1aB1bA4IV exhibited solubility similar to that of A5A4B3 rather than of WT A1aB1b (Chapter I). A1aB1b α and A1aB1bNeg exhibited similar solubility profiles to each other at low and high ionic strengths, and low solubility at low pHs at both ionic strengths. A1aB1bPos exhibited unpredicted solubility profile and mostly insoluble in all pHs at $\mu = 0.5$. The solubility of this mutant was never beyond 40%. At $\mu = 0.08$, the solubility of this mutant increased when the pH increased and reached the level of 70%. At $\mu = 0.5$, the solubility of A2B1a α' and A2B1a α were similar to that of the WT A2B1a. At $\mu = 0.08$,

they showed clear isoelectric precipitation at pH 4.0 to 5.2 (Figure 7B). These phenomena might be due to combinations of the effects of (i) neutralization by electrostatic interaction of positively charged residues at the C-terminus with the negatively charged residues at disordered regions II and IV, (ii) change of electrostatic potential of charged residues by pH, (iii) neutralization by salt, and (iv) increase of hydrophobic interaction by salt.

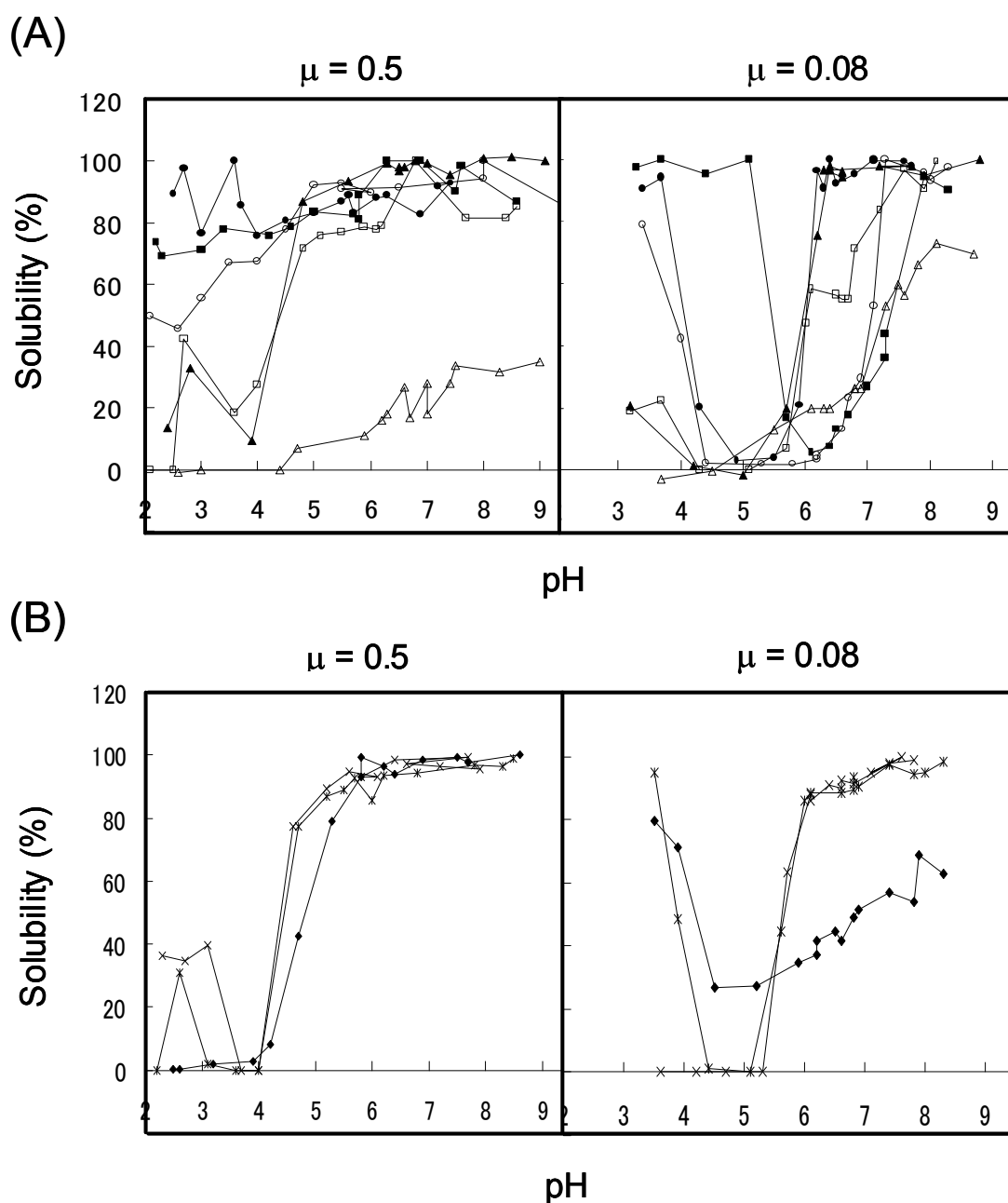


Figure 7. pH dependence of the solubility of proglycinins and their mutants. (A), A1aB1b and its mutants; (B) A2B1a and its mutants. ■, proglycinin A1aB1b WT; ●, A1aB1b α' ; □, A1aB1b α ; ○, A1aB1bA4IV; ▲, A1aB1bNeg; △, A1aB1bPos; ◆, proglycinin A2B1a WT; ×, A2B1a α' ; *, A2B1a α .

Emulsifying Property

The emulsifying property of WT and mutants was studied at high ($\mu = 0.5$) and low ($\mu = 0.08$) ionic strengths, pH 7.6 (Figures 8 and 9). The investigation was based on two criteria: emulsifying ability and emulsion stability. The emulsifying ability (Figure 8) of the proteins was analyzed by measuring the particle size distribution of the emulsion samples using light scattering instrument. At $\mu = 0.5$, the emulsions of all mutant A1aB1bs were similar in average particle size to that of the wild type A1aB1b (3.9 μm) except that of A1aB1b α' (2.4 μm) which was much better. At $\mu = 0.08$, the emulsion properties of all the mutants were better than that of the WT A1aB1b. For A2B1a α' and A2B1a α , their emulsion properties were much better than WT A2B1a at high and low ionic strengths. The orders of emulsifying ability of all the samples at $\mu = 0.5$ and 0.08 were A1aB1b $\alpha' > \text{A2B1a}\alpha' > \text{A1aB1bA4IV} \geq \text{A1aB1b}\alpha > \text{A1aB1b} > \text{A2B1a}\alpha = \text{A1aB1bPos} > \text{A1aB1bNeg} > \text{A2B1a}$ and A1aB1b $\alpha = \text{A1aB1bPos} \geq \text{A1aB1b}\alpha' > \text{A2B1a}\alpha \geq \text{A2B1a}\alpha' > \text{A1aB1bA4IV} \geq \text{A1aB1bNeg} \geq \text{A2B1a} > \text{A1aB1b}$, respectively. From these results the author suggests that the introduction of hydrophilic residues especially positively charged residues improved emulsion ability of the proglycinin mutants at low ionic strength.

The stability of the emulsion of the samples was analyzed by keeping the emulsion samples at room temperature without agitation and was visually observed after 1/24, 20/24, 2, 5, 7 and 14 days (Figure 9). Emulsion stabilities of all mutants were better than that of the wild type at both ionic strengths after 20 h and they were ordered A1aB1b $\alpha \geq \text{A2B1a}\alpha > \text{A1aB1b}\alpha' \geq \text{A2B1a}\alpha' \gg \text{A1aB1bPos} > \text{A1aB1bA4IV} \geq \text{A1aB1bNeg} > \text{A2B1a} > \text{A1aB1b}$ (Figure 9A). Among all mutants in this study, the mutants by extension of α' and α extension regions to the C-terminal of proglycinins exhibited best emulsion stabilities. The emulsion stabilities of A1aB1b α' and A1aB1b α were similar to that of A2B1a α' and A2B1a α in a later observation (data not shown). For further investigation, the A1aB1b α and A1aB1b α' emulsions were kept at room temperature and time dependence of their emulsion stabilities were recorded. It was found that the emulsions of the mutants at $\mu = 0.08$ were better than that at $\mu = 0.5$. At $\mu = 0.5$, the phase separation of

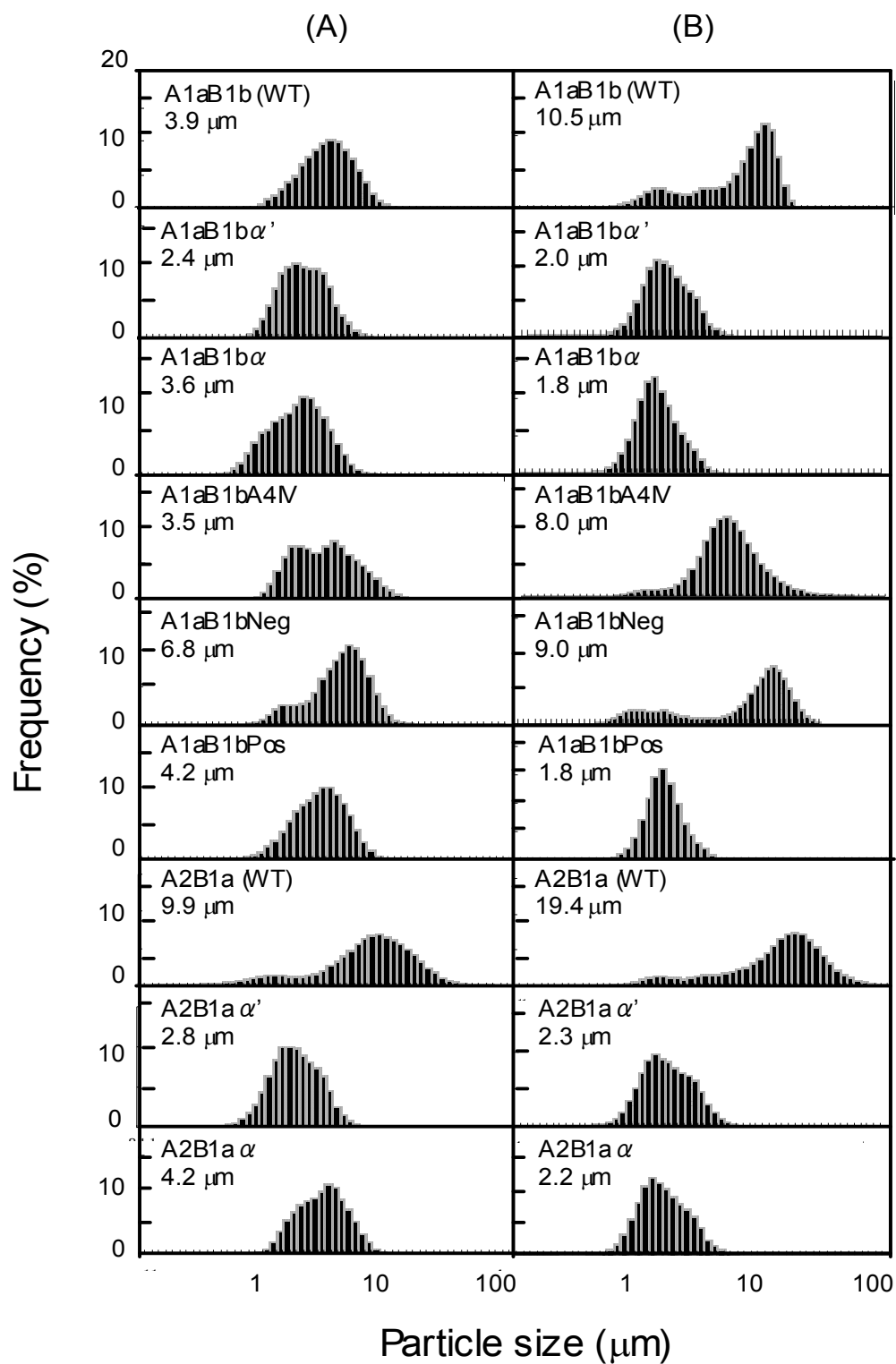
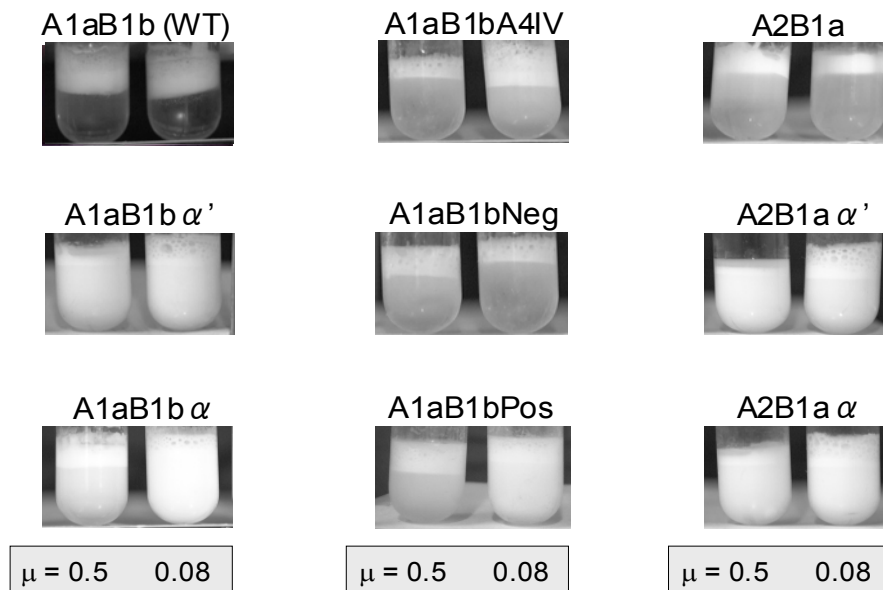


Figure 8. Particle size distributions of emulsion of progycinins and mutants at ionic strength 0.5 (A) and 0.08 (B).

(A)



(B)

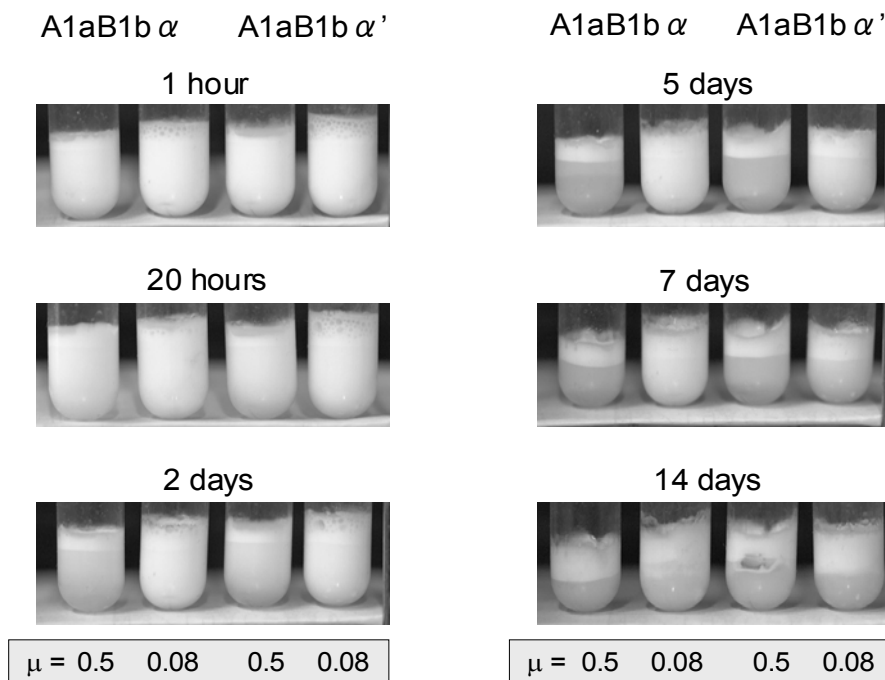


Figure 9. Emulsion stability. (A) Emulsion formed using proglycinins and their mutants as emulsifier after 20 h. (B) Time dependence of emulsion stability of A1aB1b α' and A1aB1b α at room temperature.

these emulsions was slightly observed after 2 days and clearly observed after 5 days (Figure 9B). However, there was no complete separation of cream and water was observed even after 28 days (data not shown). At $\mu = 0.08$, the phase separation of A1aB1b α' started only after 5 days and phase separation of A1aB1b α started only after 14 days. These indicated that α' and especially α extension polypeptides (extension regions) are good polypeptides for the improvement of emulsifying ability and emulsion stability of soybean proglycinins.

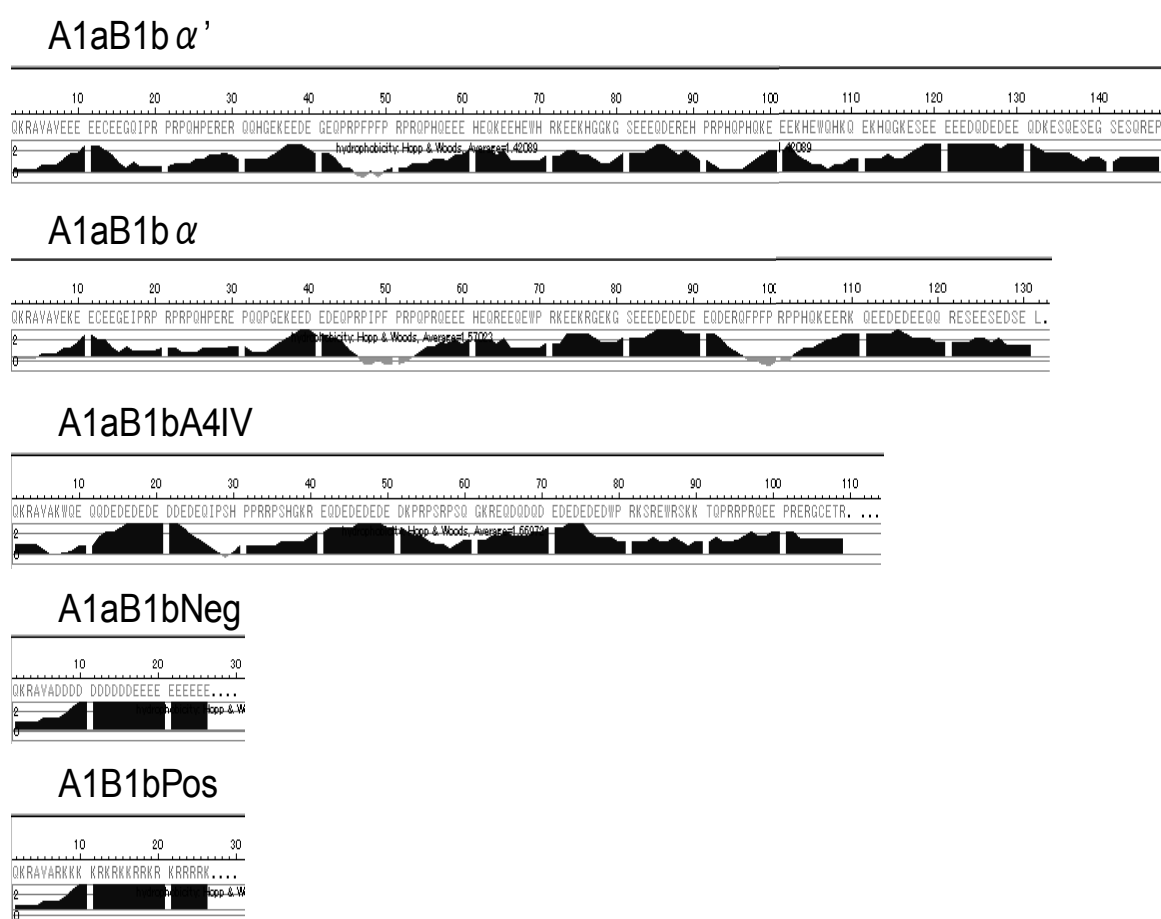


Figure 10. Hydrophobicity profiles of newly C-termini of mutants (disordered region V of A1aB1b and introduced peptides).

A1aB1b α and A2B1a α exhibited similar emulsion stability to each other and their stabilities were better than those of A1aB1b α' , A2B1a α' and A1aB1bA4IV. Tandang et al. [12] found that while changing the position of α' extension region from the N- to the C-terminus of the α' subunit resulted in better emulsion stability of the protein. It is observed that the first 20 amino acids of the α' extension N-terminus of the region were less hydrophilicity than that of the last 20 amino acids of the α' extension C-terminus. Therefore, the emulsion stability of the mutants might depend on the charged residues of the last 20 amino acids at the C-terminus of the polypeptides. It is noted that hydrophilic profile of 20 amino acids of the C-terminus of the polypeptide in this study was $\alpha > \alpha' > A4IV$ and the emulsifying property was also $A1aB1b\alpha > A1aB1b\alpha' > A1aB1bA4IV$. These findings proved the assumption that the hydrophilicity of the last 20 amino acids of the polypeptide is important for the improvement of emulsion stability. A1aB1bPos and A1aB1bNeg although had higher hydrophilic composition than that of the others, they provided emulsion stability lesser than that of A1aB1b α' and A2B1a α' . Therefore, it might be not only the distribution of the hydrophilicity of the 20 amino acid residues at the end of C-terminus but also the length of the introduced polypeptide that contributed to the emulsifying property of the protein.

For the detailed observation, hydrophobic profiles of C-termini (disordered region V of A1aB1b plus introduced peptides) of mutant A1aB1bs were analyzed by DNAsis program (hydrophobic profile by Hopp and Woods) (Figure 10). It was found that all the C-termini of mutants (newly disordered region V) were rich in hydrophilic residues. The hydrophilic composition of the C-termini was 78.2, 73.2, 78.2, 88.4, and 88.4% for A1aB1b α' , A1aB1b α , A1aB1bA4IV, A1aB1bNeg and A1aB1bPos, respectively. Among all mutant C-termini, A1aB1b α C-terminus had a bit lower hydrophilic composition than that of the others but it had strong and special hydrophobic regions corresponding to the amino acid residues number 1–7, 45–55 and 97–103 of A1aB1b α C-terminus (Figure 10). Residues 1–7, 44–52 and 91–96 of A1aB1b α' C-terminus were also hydrophobic regions of A1aB1b α' C-terminus but their hydrophobicities were poorer than that of the A1aB1b α C-terminus. A1aB1bA4IV C-terminus had hydrophobic regions at residues 4–8 and 27–32 and they showed poorer hydrophobicity than that of the A1aB1b α' C-terminus.

Among all C-termini, A1aB1b α C-terminus is best in hydrophobic/hydrophilic profile: three pairs of a strong hydrophobic patched region and a strong hydrophilic region. It was reported that in an oil-water emulsion, a protein with high hydrophobic surface would prefer being with oil and a protein with high hydrophilic surface would prefer being with water. A protein with high flexible surface and good balance of hydrophobic/hydrophilic surface was reported to be a good protein for good emulsifying property (Graham and Phillips, 1976; Nakai et al., 1980, 1983, 1986; Damodaran, 1997). In fact, A1aB1b α that has this criterion exhibited good emulsifying property than the other mutants. Therefore, hydrophobic/hydrophilic profile of the introduced polypeptide plays an important role for emulsifying property of mutant.

From all of the observations, conclusion could be made that the emulsifying property, especially emulsion stability, of proglycinin mutants depends on (i) the hydrophilicity of the 20 amino acid residues at the end of C-terminus, (ii) the length of the introduced polypeptide to the C-terminus, (iii) the distribution of the hydrophilic/hydrophobic profile of the polypeptide after introduction to the C-terminus of proglycinins and (iv) other factor such as ionic strength of the sample. Among the introduced polypeptides and oligopeptides in this study α extension region was a very good polypeptide that had the above three properties.

The information obtained from the introduction of polypeptide and oligopeptide to the C-terminus of proglycinins would be very useful for improvement of emulsifying property of food proteins.

References

- [1] Graham, D. E. and Phillips, M. C. (1976). The conformation of proteins at interfaces and their role in stabilizing emulsions. In *Theory and practice of emulsion technology*; Smith, A. L., Ed.; Academic Press: London. pp. 75–93.
- [2] Phillips, M. C. (1981). Protein conformation at liquid interfaces and its role in stabilizing emulsions and foams. *Food Technol.* **35**, 50–59.
- [3] Wagner, J. R. and Guéguen, J. (1999). Surface functional properties of native, acid-treated, and reduced soy glycinin. 2. Emulsifying properties. *J. Agric. Food Chem.* **47**, 2181–2187.
- [4] Maruyama, N., Sato, R., Wada, Y., Matsumura, Y., Goto, H., Okuda, E., Nakagawa, S. and Utsumi, S. (1999). Structure-physicochemical function relationships of soybean β -conglycinin constituent subunits. *J. Agric. Food Chem.* **47**, 5278–5284.
- [5] Liu, M., Lee, D. S. and Damodaran, S. (1999). Emulsifying properties of acidic subunits of soy 11S globulin. *J. Agric. Food Chem.* **47**, 4970–4975.
- [6] Maruyama, N., Mohamad Ramlan, M. S., Takahashi, K., Yagasaki, K., Goto, H., Hontani, N., Nakagawa, S. and Utsumi, S. (2002). Structure-physicochemical function relationships of soybean β -conglycinin heterotrimers. *J. Agric. Food Chem.* **50**, 4323–4326.
- [7] Kinsella, J. E. (1982). In *Food Proteins*; Fox, P., Condon, J. J., Eds.; Elsevier Applied Sci. Pub.: London, pp. 51.
- [8] Kinsella, J. E. (1979). Functional properties of soy proteins. *J. Am. Oil Chem. Soc.* **56**, 242–258.
- [9] Wagner, J. R. and Guéguen, J. (1995). Effects of dissociation, deamidation, and reducing treatment on structural and surface active properties of soy glycinin. *J. Agric. Food Chem.* **43**, 1993–2000.
- [10] Kim, S. H. and Kinsella, J. E. (1987). Surface active properties of food proteins: Effects of reduction of disulfide bonds on film properties and foam stability of glycinin. *J. Food Sci.* **52**, 128–131.
- [11] Maruyama, N., Katsube, T., Wada, Y., Oh, M. H., Barba de la Rosa, A. P., Okuda, E., Nakagawa, S. and Utsumi, S. (1998). The roles of the N-linked glycans and extension regions of soybean β -conglycinin in folding, assembly and structural features. *Eur. J. Biochem.* **258**, 854–862.
- [12] Tandang, M. R. G., Atsuta, N., Maruyama, N., Adachi, M. and Utsumi, S. (2005). Evaluation of the solubility and emulsifying property of soybean proglycinin and rapeseed procruciferin in relation to structure modified by protein engineering. *J. Agric. Food Chem.* **53**, 8736–8744.
- [13] Katsube, T., Kurisaka, N., Ogawa, M., Maruyama, N., Ohtsuka, R., Utsumi, S. and Takaiwa, F. (1999). Accumulation of soybean glycinin and its assembly with the glutelins in rice. *Plant Physiol.* **120**, 1063–1073.
- [14] Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.
- [15] Kim, C. S., Kamiya, S., Kanamori, J., Utsumi, S. and Kito, M. (1990). High level expression, purification and functional properties of soybean proglycinin from *Echerichia coli*. *Agri. Biol. Chem.* **54**, 1543–1550.
- [16] Burnette, W. H. (1981). Western blotting: electrophoretic transfer of proteins from SDS-polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radionated protein. *Anal. Biochem.* **112**, 95–203.
- [17] Green, A. A. and Hughes, W. L. (1955). *Methods in Enzymology*; Colowick, S. P., Kaplan, N. O., Eds.; Academic Press: New York, Vol. 1, p 67.
- [18] Adachi, M., Takenaka, Y., Gidamis, A. B., Mikami, B. and Utsumi, S. (2001). Crystal structure of soybean proglycinin A1aB1b homotrimer. *J. Mol. Biol.* **305**, 291–305.

- [19] Nakai, S. (1983). Structure-function relationships of food proteins with an emphasis on the importance of protein hydrophobicity. *J. Agric. Food Chem.* **31**, 676–683.
- [20] Nakai, S. and Li-Chan, E. (1988). Hydrophobicity-functionality relationship of food proteins. *Hydrophobic Interactions in Food Systems*; CRC Press: Boca Raton, FL, pp 23–41.
- [21] Kato, A. and Nakai, S. (1980). Hydrophobicity determined by a fluorescence probe method and its correlation with surface properties of proteins. *Biochim. Biophys. Acta* **624**, 13–20.
- [22] Hayakawa, S. and Nakai, S. (1985). Relationships of hydrophobicity and net charge to the solubility of milk and soy proteins *J. Food Sci.*, **50**, 486–491.
- [23] Nakai, S., Ho, L., Helbig, N., Kato, A. and Tung, M. A. (1980). Relationship between hydrophobicity and emulsifying properties of some plant proteins. *Can. Inst. Food Sci. Technol. J.* **13**, 23–27.
- [24] Nakai, S. (1983). Structure-function relationship of food proteins with emphasis on the importance of protein hydrophobicity. *J. Agric. Food Chem.* **31**, 676–683.
- [25] Nakai, S., Li Can, E. and Hayakawa, S. (1986). Contribution of protein hydrophobicity to its functionality. *Nahrung* **30**, 327–336.
- [26] Damodaran, S. (1997). Protein-stabilized forms and emulsions. In *Food Proteins and Their Applications*; Damodaran, S., Paraf, A., Eds.; Marcel Dekker: New York, pp. 57–110.

Chapter IV

Design of Genetically Modified Soybean Proglycinin A1aB1b with Multiple Copies of Bioactive Peptide Sequences

Some bioactive peptides derived from food proteins have beneficial effects on human health. For the past three or four decades, various studies on bioactive peptides in amino acid sequences of natural proteins such as cow milk [1–3], egg white [4], fish [5,6], soybean [7,8], rapeseed [9], spinach [10], mushroom [11,12], and many other sources have been reported. These protein-derived bioactive peptides are inactive within the sequences of the parent proteins but can be released by enzymatic proteolysis during gastrointestinal digestion. Once liberated in the body, bioactive peptides may act as regulatory compounds with hormone-like activity. Thus, these peptides represent potential health enhancing components for food and pharmaceutical applications. However, the amount of derived bioactive peptide is low or does not exist in some types of natural proteins. The appropriate bioactive peptides can, however, be introduced into food proteins and/or concentrated for the purpose of treatment of a disease or disorder. With this aim, attempts have been made to introduce bioactive peptides derived from a protein source to other proteins and to improve the level of bioactive peptide content in the protein itself [13–18]. However, some of the expressed mutant proteins containing the bioactive peptide were not soluble in *Escherichia coli* [13,17,18], suggesting that they could not form proper conformations.

The author's ultimate purpose is to develop a crop that produces modified proteins containing useful bioactive peptides in as high amount as possible. To attain this aim, the modified proteins should form correct conformations. Proteins which have misfolded might be degraded by proteinases that are present in vacuoles [19–21], hence these misfolded proteins will not be able to accumulate as storage proteins. Therefore, after introduction of mutations, the resulting modified proteins should be evaluated if they can form the correct conformation or not. Earlier to evaluate the correct conformation of

proglycinin, Utsumi's groups established three criteria: (1) high expression in *E. coli* (> 10% of *E. coli* total protein), (2) solubility comparable with that of globulins, and (3) self-assembly into trimers. They have reported developing modified proglycinins that satisfied these three criteria and which were shown to accumulate in the protein storage vacuole without difficulty [22–24].

Among seed storage proteins, soybean protein is a good model for site-directed mutation since the protein content of soybean is the highest among major crops. The modified protein would be available for various uses since the utilization of soybean protein isolate, which is generated as a by-product of soybean oil, has been well established. Among the glycinin five subunits A1aB1b, A1bB2, A2B1a, A3B4 and A5A4B3, the A1aB1b subunit was selected for the introduction of a bioactive peptide sequence. This is due to the following reasons: (1) among the five subunits of glycinin, it is highly expressed as soluble protein in *E. coli* (Chapter II); (2) it has a molecular chaperone-like function for the other subunits [25]; and (3) it has bile acid-binding ability directed by the sequence VAWWMY [13] that could inhibit the reabsorption of bile acid in the ileum or decrease the micellar solubility of cholesterol in small intestinal epithelial cells, resulting in decreasing the blood cholesterol level [26–29].

A primary consideration in this study is the choice of the site on A1aB1b which should be mutated. All five variable regions (disordered regions) of A1aB1b play a minor role in proper folding of the glycinin [30]. Therefore, the introduction of a bioactive peptide into any of these variable regions of this polypeptide would not cause a change in its folding property. In fact, mutations in the variable region IV of A1aB1b did not affect its correct folding [31,32]. On the other hand, Takenaka et al. [18] observed that a small change (mutation) of a few amino acids in the conserved region of A1aB1b resulted in the localization of the expressed protein in the inclusion body, indicating that the modified protein had the wrong conformation and was insoluble.

The short peptide IIAEK is a hypocholesterolemic peptide derived from a food protein, bovine β -lactoglobulin. It contains nearly equal amounts of hydrophobic and charged residues, and has no peptic or tryptic digestive site. The hypocholesterolemic activity of

IIAEK has been reported to be greater compared to that of a medicine, β -sitosterol, in rats [3]. Thus, IIAEK is a good model peptide for our present study. Multiple repeats of the IIAEK sequence were introduced into the five disordered regions of soybean proglycinin A1aB1b. The modified proteins were expressed in *E. coli* to assess their folding abilities. The expressed A1aB1b mutants containing high amounts of IIAEK with high folding abilities were purified and the yield of IIAEK released after in vitro digestion of these mutants with trypsin was measured.

Materials and Methods

Bacterial Strain, Medium and Plasmid

E. coli BL21(DE3), HMS174(DE3), AD494(DE3) and Origami(DE3) (Novagen) were used as hosts for expression and grown on Luria-Bertani (LB) culture medium containing 0.17 and 0.5 M NaCl, pH 7.5. The expression plasmid used for mutations was pEA1aB1b [22], which contains the cDNA encoding the proglycinin A1aB1b between the *Nco*I and *Bam*HI sites of the expression vector pET-21d (Novagen).

Insertion and Substitution Mutagenesis

The expression plasmid pEA1aB1b was used as a template for mutation in the five disordered regions of A1aB1b (Fig. 1A). To determine whether keeping the size of the subunit after mutation close to its original size was required for proper folding and higher expression of the protein, both substitution and insertion mutagenesis were attempted. Four IIAEK sequences were substituted for 4, 14 and 15 amino acids in the middle of the disordered region I (I-4), II (II-4), and III (III-4), respectively. The replaced residues were 4-7, 94-107, 181-195 of A1aB1b (Fig. 1B). In the case of the disordered region IV, the insertion of four IIAEK sequences was made at two positions, at the start (IV-S-4) and/or from the last (IV-L-4) nine amino acids of the disordered region, corresponding to the insertion between residue numbers 257-258 and 287-288, respectively. For the mutation in the disordered region V, four IIAEK sequences were added to the C-terminus (residue 476, V-4). Additional insertion of one, three and six IIAEK sequences in IV-L-4 or four IIAEK sequences in V-4 were done to obtain IV-L-5, IV-L-7 and IV-L-10 or V-8,

respectively (Fig. 2). The DNAs were obtained by annealing of 5'-(ATC ATC GCC GAG AAA)_n-3' and 5'-(TTT CTC GGC GAT GAT)_n-3' (n, number of appropriate repeated

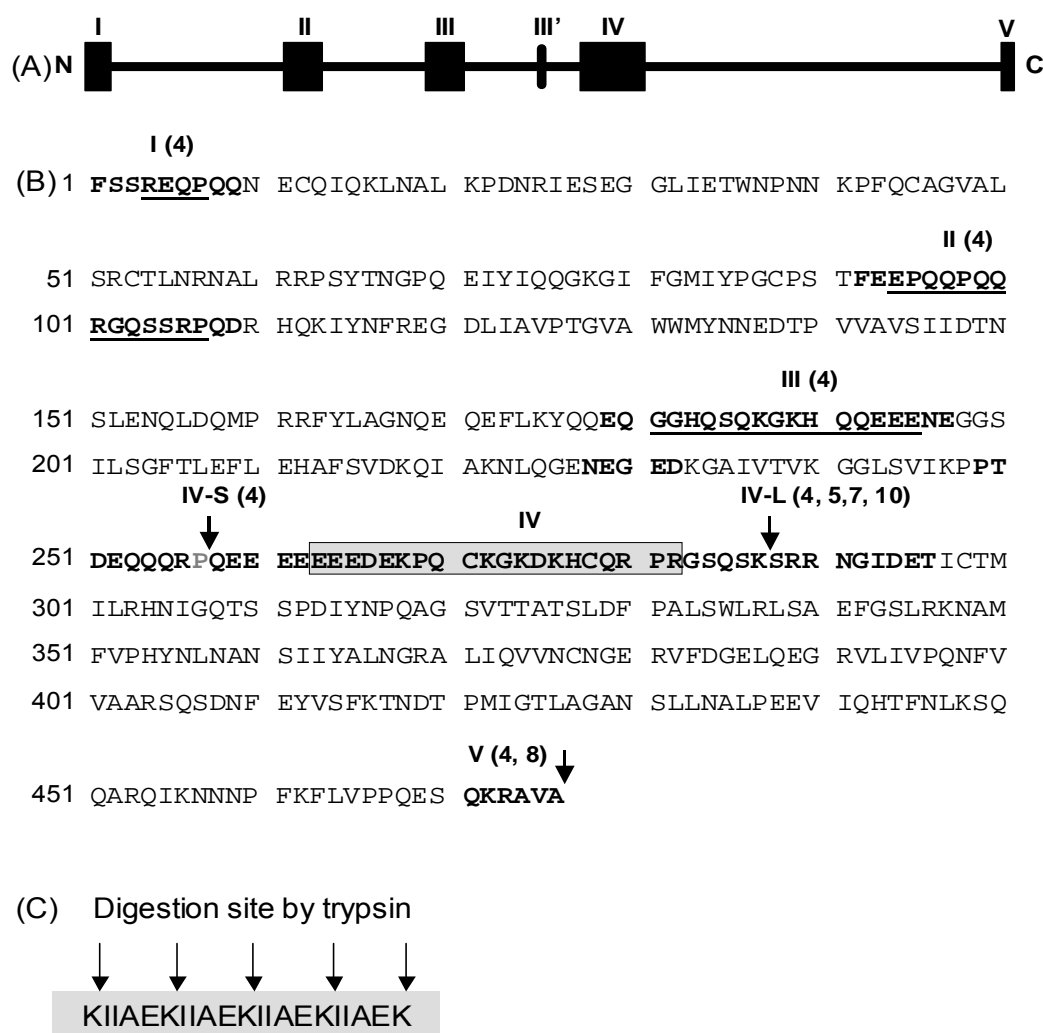


Figure 1. The schematic representation and sequence of proglycinin A1aB1b with disordered and mutated regions. (A) Schematic representation of disordered regions of A1aB1b. The black boxes and stick indicate the disordered and ordered regions, respectively. Each disordered regions were named I, II, III, III', IV and V. (B) The sequence of A1aB1b. Disordered regions were marked in bold and mutated sequences and sites were marked by underline or arrow for IIAEK substitution or insertion, respectively. The Arabic numeral in the bracket represents the number of IIAEK sequence introduced for the mutation. The letters in the gray box indicate the amino acids deleted for the mutant IV-4+4-cut. The gray letter shows the amino acid changed to glycine. (C) Sequence of four IIAEK model peptides used for the mutation.

sequences encoding IIAEK) at 67 °C for 15 min after heating at 95 °C for 4 min, followed by slow cooling down to room temperature. To get eight IIAEK sequences into the disordered region IV (IV-S-4+L-4), the combination of the two-insertions was made by using primers for IV-S-4 and IV-L-4 expression plasmid as template for the PCR. IV-4+4-cut was made by PCR using IV-S-4+L-4 as template, resulting in a deletion of 20 amino acids corresponding to residues 263–282 in the middle region between the two insertion sites. The combinations of various mutations in the disordered regions were made by substitution of the mutation regions for the wild type regions via restriction enzyme sites: *Sma*I, *Avr*II, *Nsi*I and *Bsu*36I of A1aB1b, and *Xba*I and *Bam*HI of pET-21d.

For high releasing efficiency of IIAEK from the modified A1aB1b subunit in the intestine after ingestion of the protein, all mutants except the insertion mutants IV-L-4, IV-L-5, IV-L-7 and IV-L-10 where the insertion position 287 was lysine, a lysine residue was added to the first IIAEK sequence resulting into KIIAEK. This addition allows the IIAEK sequences to be released by trypsin. The proline residue at position 257 which is encoded by codons CCC is undesirable for digestion by trypsin, and was thus changed to glycine, a small amino acid encoded by GGC codons, which would not prevent digestion with trypsin.

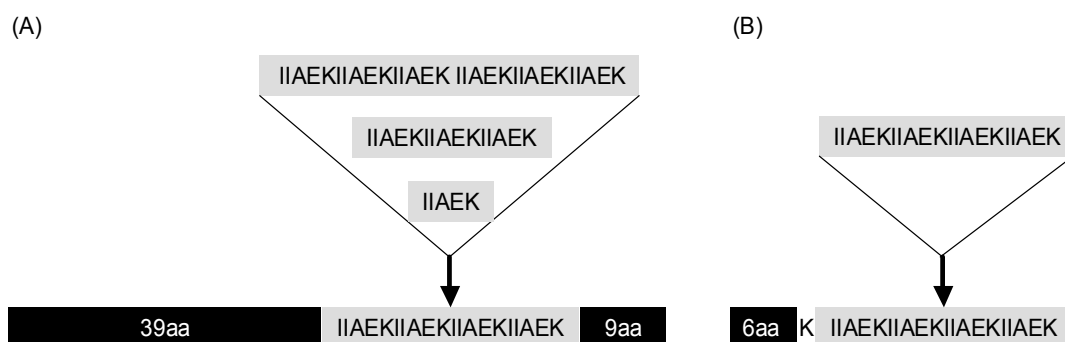


Figure 2. Additional insertion of IIAEK sequences to the mutated variable region IV and V. (A) Additional insertion to mutated variable region IV (IV-L-4) with one, three or six IIAEK sequences; (B) Additional insertion to mutated variable region V (V-4) with four IIAEK sequences.

The primers used to conduct the mutations were as follows: for I-4, 5'-**ATT ATT GCG GAA AAA ATT ATT GCG GAA AAA CAG CAA AAC GAG TGC CAG ATC CAA**-3' and 5'-**TTT CTC GGC GAT GAT TTT CTC GGC GAT GAT TTT** GGA ACT GAA CAT GGT ATA TCT CCT-3'; II-4, 5'-**ATT ATT GCG GAA AAA ATT ATT GCG GAA AAA CAA GAC CGT CAC CAG AAT ATC TAT**-3' and 5'-**TTT CTC GGC GAT GAT TTT CTC GGC GAT GAT TTT** TTC AAA TGT GCT AGG ACA ACC CGG-3'; III-4, **ATT ATT GCG GAA AAA ATT ATT GCG GAA AAA AAC GAA GGA GGC AGC ATA TTG AGT**-3' and 5'-**TTT CTC GGC GAT GAT TTT CTC GGC GAT GAT TTT** TTG CTC TTG CTG ATA TTT TAG AAA-3'; IV-L-4, 5'-**ATT ATT GCG GAA AAA ATT ATT GCG GAA AAA AGC AGA AGA AAT GGC ATT GAC GAG**-3' and 5'-**TTT CTC GGC GAT GAT TTT CTC GGC GAT GAT TTT** GCT TTG GCT TCC TCG GGG GCG-3'; IV-S-4, 5'-**ATC ATT GCA GAG AAG ATC ATT GCA GAG AAG CAG GAA GAG GAA GAA GAA GAG**-3' and 5'-**TTT TTC AGC GAT AAT TTT TTC AGC GAT AAT TTT GCC** TCT TTG TTG CTG CTC GTC CGT-3'; IV-S-4+L-4, primers used were same for IV-S-4; IV-4+4-cut, 5'-GGA AGC CAA AGC AAA ATC-3' and 5'-TTC TTC CTC TTC CTG CTT-3'; V-4, 5'- **ATT ATT GCG GAA AAA ATT ATT GCG GAA AAA TAG AAT TCC GGA TCC GAA TTC GAG**-3' and 5'- **TTT CTC GGC GAT GAT TTT CTC GGC GAT GAT TTT** AGC CAC AGC TCT CTT CTG AGA CTC-3' (thick, italic, and underlined letters indicate the sequences encoding for IIAEK, the codon for additional lysine for excision with trypsin and the nucleotide change from cytidine, respectively).

The mutated DNA sequences were confirmed by the dideoxy chain-termination method of Sanger et al. [33] using a DNA sequencer (Applied Biosystems, Model 3100 Avant) and an ABI dye terminator cycle sequencing kit with *Amplitaq* polymerase FS (Perkin Elmer, Applied Biosystems).

Expression and Solubility of Mutants Containing IIAEK Sequences in *E. coli*

The mutated plasmids were transformed into expression hosts BL21(DE3), HMS174(DE3), AD494(DE3) and Origami(DE3). Forty µl of fully-grown cultures were added to the fresh 4 ml LB media containing 0.17 and 0.5 M NaCl, and cultured at 37 °C.

At $A_{600} = 0.3\text{--}0.6$, isopropyl-1-thio- β -D-galactopyranoside (IPTG) was added to a final concentration of 1 mM to induce the expression of the mutants. Following induction, cultivation was continued at 20 °C for 40 h. The induced cells were harvested by centrifugation and stored at -20 °C until use. Cells were resuspended (15 mg cells/ml buffer) in buffer A (35 mM potassium phosphate, pH 7.6, 0.5 M NaCl, 10 mM 2-mercaptoethanol, 1 mM EDTA, 0.1 mM *p*-amidinophenyl-methylsulfonyl fluoride (APMSF), 1 μ g/ml leupeptin, 1 μ g/ml pepstatine A and 0.02% NaN₃) and disrupted by sonication on ice. The insoluble fraction was separated from the soluble fraction by centrifugation at 14,000 rpm (17,800 $\times g$) for 10 min. The total, soluble and insoluble fractions were analyzed by SDS–PAGE [34] using 11% acrylamide [35].

Purification of Mutants Containing Eight IIAEK Sequences

All procedures were carried out at 4 °C and centrifugations were at 10,000 rpm (9100 $\times g$) for 20 min unless otherwise stated. The basic buffer for all purification steps was buffer A. Ammonium sulfate fractionation followed the procedure of Green and Hughes [36].

The cells were resuspended in buffer A (40.0 mg cells/ml buffer), and disrupted by sonication. Insoluble materials were separated from the soluble fraction by centrifugation. The soluble crude extract was adjusted to 35% (for II-4+III-4) and 15% (for IV-4+V-4) saturation of ammonium sulfate and stirred for 1 h on ice. After centrifugation, the supernatant was applied on a Toyopearl (Butyl-650M) (TOSOH) column (2.6 cm x 20 cm) equilibrated with buffer A containing 35% and 15% ammonium sulfate for II-4+III-4 and IV-4+V-4, respectively. Elution was carried out with a linear gradient (800 ml) from 35 and 15 to 0% for II-4+III-4 and IV-4+V-4, respectively. Fractions containing expressed proteins were concentrated to 2 ml by Centriprep YM-10 (Millipore), and then the concentrated samples were subjected to gel filtration on a Hi-Prep 26/60 Sephacryl S-300 HR (Pharmacia Biotech), using buffer A as a mobile phase. For II-4+III-4, fractions containing the expressed protein from the gel filtration column were dialyzed against buffer A with 0.15 M NaCl instead of 0.5 M, and then the dialyzed sample was applied on Mono Q HR 10/10 column (Pharmacia Biotech) equilibrated with the same buffer. The

elution was performed with a linear gradient from 0.15 to 0.4 M NaCl in buffer A over a period of 80 min at 1 ml/min.

Protein Measurement

The amount of protein in the samples was determined using a Protein Assay Rapid Kit (Wako) with bovine serum albumin (BSA) as a standard.

Analysis of Self-Assembly Into Trimers

Self-assembly of individual mutants was analyzed using Hi-Prep 16/60 Sephacryl S-300 HR column (Pharmacia Biotech). The experimental conditions were the same as described previously [37] except for the use of S-300 instead of S-200 particle size.

In Vitro Digestion of Mutants Containing Eight IIAEK Sequences

Purified mutants containing eight IIAEK sequences (II-4+III-4 and IV-4+V-4) were dialyzed against milliQ water and the samples (1 mg/ml) were adjusted to pH 8.0 with 1 N NaOH, and digested with trypsin (E/S = 1/100) (w/w) for 12 h at 37 °C. The pH was checked and adjusted every hour for the first 5 h of the reaction. The reaction was stopped by adding trifluoroacetic acid (TFA) to a final concentration of 0.1% and boiled for 10 min.

Purification and Detection of IIAEK Peptide From Digestion of Mutants Containing Eight IIAEK Sequences

The digests of II-4+III-4 and IV-4+V-4 were fractionated by reverse phase HPLC on an octadecyl silica (ODS) column (Cosmosil 5C18-AR-II, 4.6 x 150 mm, Nakalai Tesque). Elution was performed with a linear gradient of acetonitrile (1%/3min), containing 0.1% TFA at a flow rate of 1 ml/min and monitored at 230 nm. Synthesized IIAEK peptide (purity > 99.7%, Sigma Genosys Japan, Hokkaido, Japan) was used as reference for the elution time. The expected fractions containing the IIAEK peptide were collected. To confirm the identity of the peptide, the samples were applied on a liquid chromatograph

mass spectrometry (Mariner, PerSeptive Biosystems) and a 492 protein sequencer (Applied Biosystems).

The amount of the purified peptide was estimated from the peak area of absorbance at 230 nm on HPLC by using the synthesized peptide as a standard. The overall yield (%) was calculated as follows: [purified peptide (mol)/theoretical peptide released after digestion (mol)] x 100.

Results

Expression and Solubility of Mutants in *E. coli*

A synthesized peptide IIAEKIIAEKIIAEK was digested *in vitro* with trypsin at pH 8.0 for 5 h. The author found that the recovery of IIAEK was nearly 100%, i.e. one mole of the peptide gave nearly three moles of IIAEK (data not shown). Thus, it was expected that such a tandem of repeated peptide with additional lysine in front of the first isoleucine will be efficiently digested by trypsin. The sequences for our mutations were designed based on this property (Fig. 1C).

Considering the balance of expression level and solubility of expressed modified A1aB1b (data not shown), the author selected the LB medium containing 0.5 M NaCl as medium and Origami (DE3) for II-4+III-4 and AD494(DE3) for the other mutants as host *E. coli*. The level of expression (%) and solubility (%) of each mutant containing a single set of four IIAEK sequences were: 25 and 25, 12 and > 80, 10 and > 80, 25 and > 80, 25 and > 80, 25 and > 80 for I-4, II-4, III-4, IV-S-4, IV-L-4 and V-4, respectively (Fig. 3, A–F). Both the solubility and expression levels of the mutants IV-S-4, IV-L-4 and V-4 were the best among these mutants. The author increased the number of tandem repeated IIAEK sequences introduced into the disordered region IV and V to get five to ten of IIAEK sequences in these disordered regions. The level of expression (%) and solubility (%) were: 20 and > 80, 12 and 40, 25 and < 5, 25 and ~ 1 for IV-L-5, IV-L-7, V-8 and IV-L-10, respectively (Fig. 3, G–J). IV-4+4 which resulted from a combination of IV-S-4 (Fig. 3D) and IV-L-4 (Fig. 3E) gave poor expression level. Deletion of 20 amino acids from a

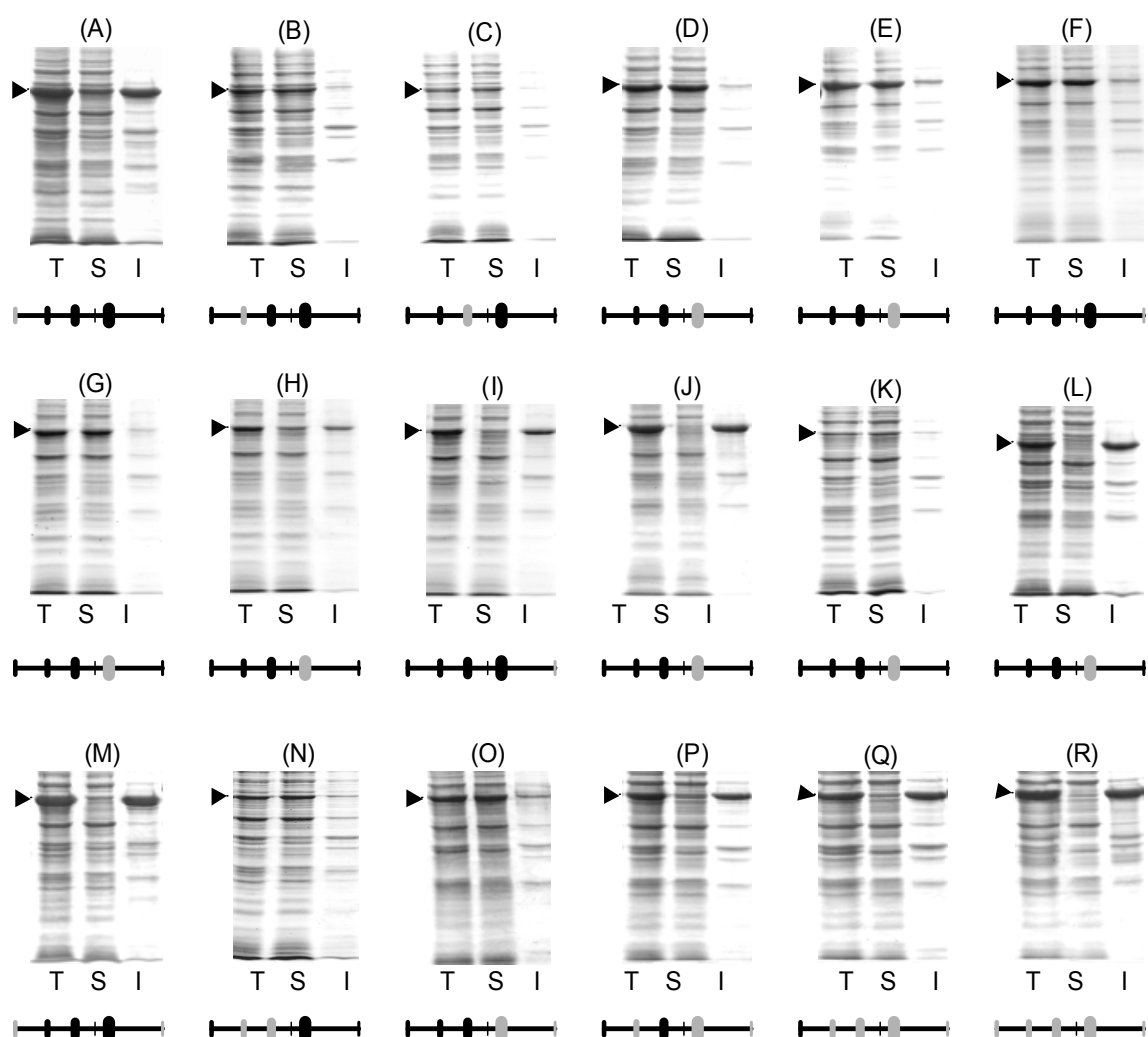


Figure 3. SDS-PAGE (11% acrylamide) profile of *E. coli* expressing proglycinin A1aB1b-containing IIAEK sequence. The arrowhead indicates position of the expressed mutant protein. T, total protein after cell extraction; S, soluble fraction; I, insoluble fraction. (A) I-4; (B) II-4; (C) III-4; (D) IV-S-4; (E) IV-L-4; (F) V-4; (G) IV-L-5; (H) IV-L-7; (I) V-8; (J) IV-L-10; (K) IV-4+4; (L) IV-4+4-cut; (M) I-4+V-4; (N) II-4+III-4; (O) IV-4+V-4; (P) II-4+IV-4+V-4; (Q) II-4+III-4+IV-4+V-4; (R) I-4+II-4+III-4+IV-4+V-4. Ball and stick represent disordered and ordered regions of A1aB1b, together with the mutated region, indicated by gray ball.

disordered region IV from this mutant resulted in IV-4+4-cut. This deletion lowered the solubility from 50 to < 2% but improved the expression level from 5 to 25% of the mutant (Fig. 3, K and L). The performance of combinations of modified disordered regions containing IIAEK sequences reduced the level of soluble expressed proteins and this level

was always lower than that of the poorest one among the originals. For example, the mutant I-4+V-4 (Fig. 3M) from a combination of I-4 (Fig. 3A) and V-4 (Fig. 3F) gave a solubility level poorer than that of I-4 the solubility of which is poorer than that of V-4. The level of expression (%) and solubility (%) of the combination mutants were: 25 and < 5, 10 and ~ 80, 20 and ~ 80, 25 and ~ 5, 25 and < 5, 25 and ~ 1 for I-4+V-4, II-4+III-4, IV-4+V-4, II-4+IV-4+V-4, II-4+III-4+IV-4+V-4, I-4+II-4+III-4+IV-4+V-4, respectively (Fig. 3, M–R). The author found that among the mutants which resulted from the combinations, only II-4+III-4 (Fig. 3N) and IV-4+V-4 (Fig. 3O) gave good solubility and expression level. The expression of mutants containing high amount of IIAEKs from many combinations remained high though the solubility became much poorer. Finally, mutants II-4+III-4 and IV-4+V-4 containing eight IIAEK sequences that gave fairly good expression and solubility levels among all combinations were selected for protein expression, purification and tryptic digestion in vitro.

Purification of II-4+III-4 and IV-4+V-4

Densitometric scanning of the SDS–PAGE electrophoregrams indicates that II-4+III-4 and IV-4+V-4 accumulated in *E. coli* at a level of around 10 and 20% of the total bacterial proteins, and that more than 80% of the recombinant proteins were soluble (Fig. 4, A and B). Immunoblotting revealed that both mutants were recognized by anti-glycinin antibody similar with the wild type recombinant A1aB1b (data not shown). The expressed proteins in the soluble fractions were first fractionated with 35 and 15% ammonium sulfate saturation for II-4+III-4 and IV-4+V-4, respectively. The soluble fractions containing the recombinant proteins were subjected to Butyl Toyopearl hydrophobic column followed by gel filtration column. The purity of mutant IV-4+V-4 was higher than 95% after gel filtration column (Fig. 4B, lane b). However, that of II-4+III-4 was only around 70% (Fig. 4A, lane b). An additional step using Mono Q column was necessary for the purification of II-4+III-4 and its purity increased to higher than 90% (Fig. 4A, lane c).

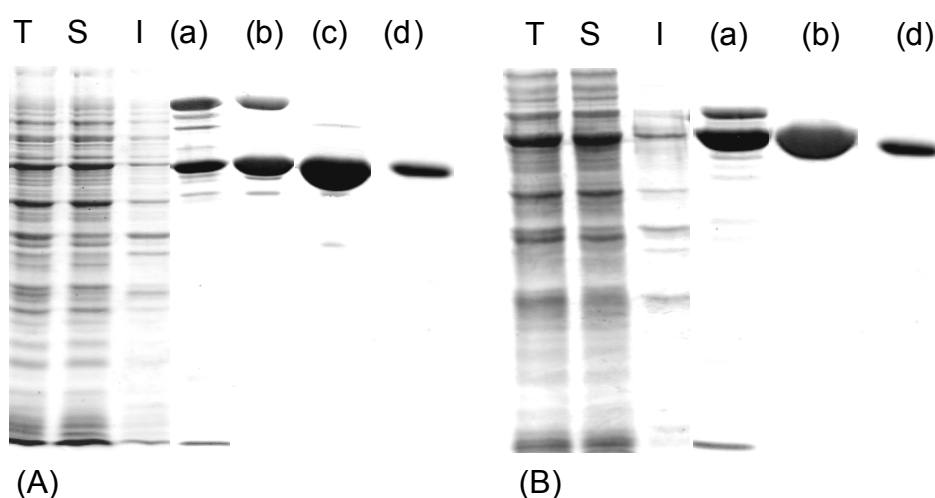


Figure 4. Expression and purification of genetically modified proglycinin A1aB1b containing eight IIAEK sequences, II-4+III-4 and IV-4+V-4. T, total protein after cell extraction; S, soluble fraction; I, insoluble fraction. (A) II-4+III-4; (B) IV-4+V-4; (a), (b) and (c) after purification by hydrophobic, gel filtration and ion exchange column chromatography, respectively; (d) purified proglycinin A1aB1b.

Self Assembly of II-4+III-4 and IV-4+V-4

It is important to analyze the self-assembly of individual mutants into trimers. Previously, Utsumi's group demonstrated that proglycinin A1aB1b self assembles into trimers [30,35,38]. Analysis by gel filtration column chromatography at pH 7.6 and ionic strength $I = 0.5$ demonstrated that the elution time of A1aB1b, II-4+III-4 and IV-4+V-4 were 123.1, 120.0, and 114.3 min, respectively (Fig. 5). These values correspond well to their trimeric molecular sizes, 160.8, 165.0, 174.9 kDa, respectively (Table 1), indicating that both II-4+III-4 and IV-4+V-4 self-assembled into trimers. Therefore, these two mutants formed the 3-D structure similar to that of A1aB1b [30].

Table 1. Molecular mass of A1aB1b and its mutants

| | Molecular mass (kDa) | |
|------------|----------------------|---------|
| | monomer | trimers |
| Wt. A1aB1b | 53.6 | 160.8 |
| II-4+III-4 | 55.0 | 165.0 |
| IV-4+V-4 | 58.3 | 174.9 |

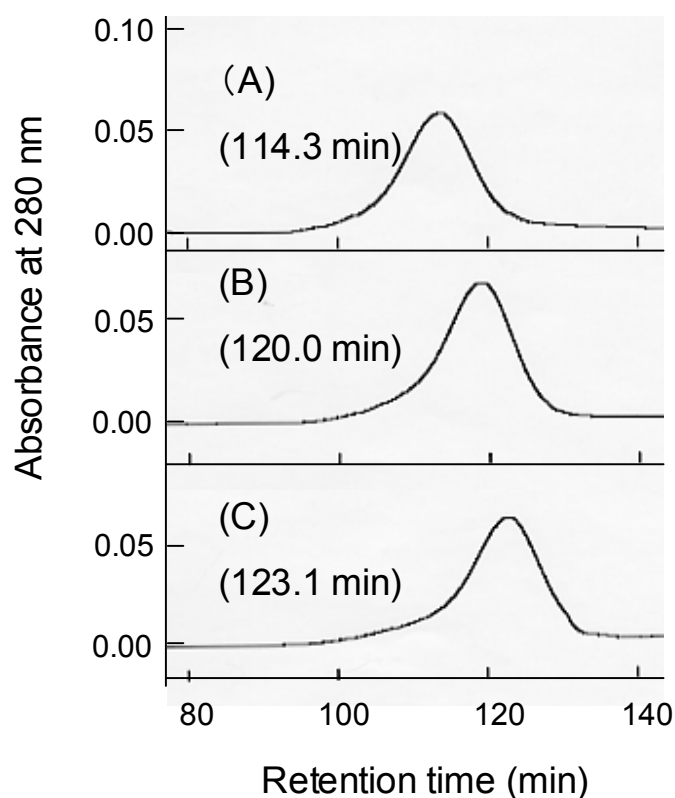


Figure 5. Gel filtration profile of proglycinin A1aB1b and its mutants. The purified proteins were subjected to chromatography using a Sephacryl S-300 HR column. (A) IV-4+V-4; (B) II-4+III-4; (C) proglycinin A1aB1b.

In Vitro Digestion of II-4+III-4 and IV-4+V-4 and Detection of IIAEK Peptide

To determine whether IIAEK peptide is released efficiently by tryptic digestion, II-4+III-4 and IV-4+V-4 were digested in vitro. After fractionation of the digests by HPLC using an ODS column, the fraction with the same retention time as that of synthetic IIAEK was analyzed by liquid chromatograph mass spectrometry, and a peak with the same molecular mass as that of IIAEK was detected (Fig. 6). For further verification, the fraction was subjected to protein sequencing. The sequence determined here agreed completely with IIAEK. The yield of IIAEK from purified protein was 77 and 90% for II-4+III-4 and IV-4+V-4, respectively.

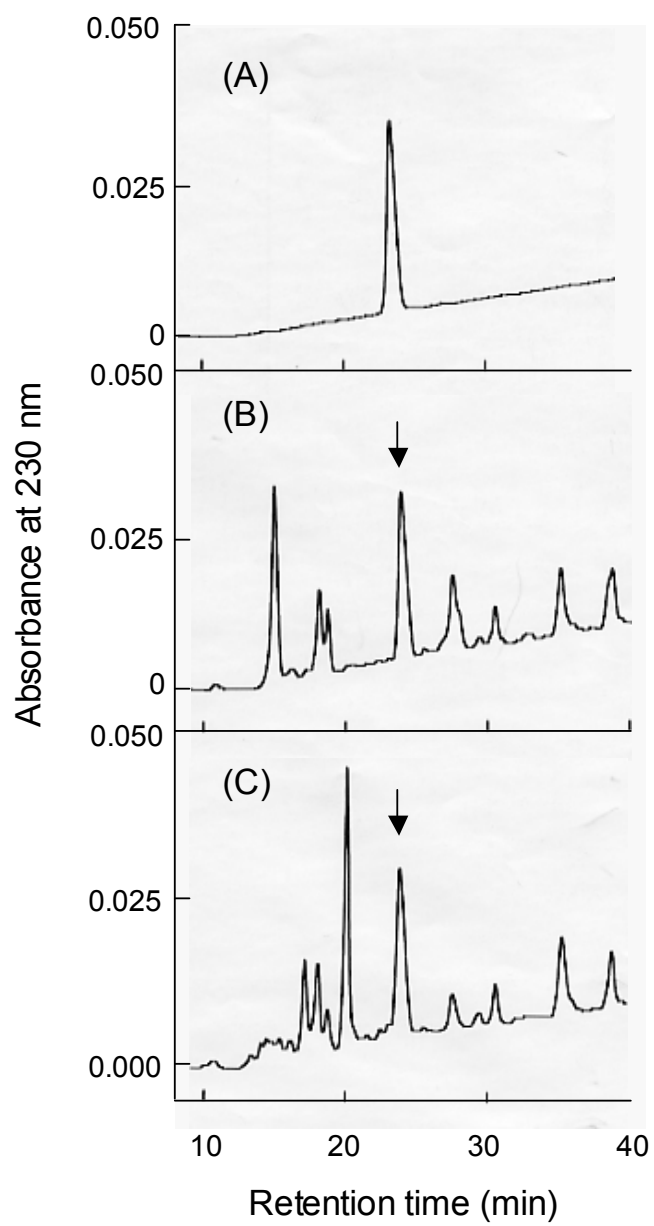


Figure 6. Purification of IIAEK from mutant tryptic hydrolysate by octadecyl silica (ODS) column. (A) standard synthetic IIAEK; (B) II-4+III-4; (C) IV-4+V-4. Arrow indicates the peak of IIAEK.

Discussion

The author introduced a sequence composed of four contiguous IIAEKs into each of the five disordered regions of proglycinin A1aB1b. Since the lengths of the disordered regions I, II and III are short, the sequence was introduced by substitution for 4, 14 and 15 amino acids. The sequence was inserted at two positions (IV-S and IV-L) of the disordered region IV and added at the C-terminus of the disordered region V, because the disordered regions IV and V are very long and located at the C-terminus, respectively. Among these mutants containing a single set of four contiguous IIAEKs, only mutant I-4 was highly insoluble. In case of I-4, 17 amino acids were inserted as a result of partial substitution, suggesting that disordered region I cannot become too long. This is consistent with Utsumi's group previous observation that proglycinin A1aB1b with the extension domain (141 amino acids) of β -conglycinin α' subunit could not accumulate in *E. coli* [39]. On the other hand, II-4, III-4, IV-S-4, IV-L-4 and V-4 exhibited high solubility, suggesting that the tolerance to modification of disordered regions II, III, IV and V was quite high. The expression of II-4 and III-4 was lower than the others, but the reason for this observation is unclear.

The author examined the effect of the length of the peptide introduced into the disordered regions IV and V. In the case of disordered region IV, she increased the repeated number of the IIAEK sequence from four to five (IV-L-5), seven (IV-L-7) and ten (IV-L-10). As the number of the inserted sequence increased, the solubility of the resulting mutant decreased, although the solubility of IV-L-5 was very high. In the case of disordered region V, she tried only V-8, which exhibited very low solubility.

When Utsumi's group introduced the extension domain of β -conglycinin α' subunit into the disordered region IV and the C-terminus of the disordered region V, the proglycinin A1aB1b accepted these modifications. IV-L-7, V-8 and IV-L-10 had extra 35, 41 and 50 amino acids which are less than the 141 amino acids of the α' extension domain. However, the introduction of seven, eight and ten IIAEK sequences apparently was not favorable to the proteins' folding into correct conformation resulting in their very low solubility. It is noted that the extension domain is very rich in charged and hydrophilic residues compared

with the multiple copies of IIAEK. Therefore, the amino acid composition of the introduced peptides is more important than their lengths to produce highly expressed soluble proteins.

Combination of single modifications resulted into lower solubility than those of the original, indicating that the undesirable effects brought about by the modifications were additive. Deletion of 20 amino acids from the disordered region IV of IV-4+4 caused complete insolubilization. The deleted region is composed of 6 negatively charged, 7 positively charged, 2 polar and 3 non-polar amino acids, indicating that this region is extremely hydrophilic. Thus, the surface hydrophilicity of IV-4+4-cut is lower than that of IV-4+4. This could be the reason why the solubility of IV-4+4-cut is much less than that of IV-4+4. Based on this, insertion mutation may be desirable for proglycinin A1aB1b than substitution mutation. In order to confirm this, modification of proglycinin A1aB1b by introduction of other peptide sequences is now in progress.

The information obtained from the introduction of IIAEK sequences into proglycinin A1aB1b would be useful in the introduction of such peptides into the other proglycinin subunits and other 11S globulin pro-forms, because they share structural characteristics with glycinin A1aB1b. However, it is to be expected that there will be differences in the acceptability of introducing various peptides among individual proglycinin subunits and among various 11S globulin pro-forms, because their folding abilities and the characteristics of disordered regions are not uniform [Chapter II,40].

References

- [1] Maruyama, S., Mitachi, H., Awaya, J., Kurono, M., Tomizuka, N. and Suzuki, H (1987). Angiotensin I-converting enzyme inhibitory activity of the C-terminal hexapeptide of α_{s1} -casein. *Agric Biol Chem.* **51**, 2557–2561.
- [2] Mullally, M. M., Meisel, H. and Gerald, F. R. J. (1997). Identification of a novel angiotensin-I-converting enzyme inhibitory peptide corresponding to a tryptic fragment of bovine β -lactoglobulin. *FEBS Lett.* **402**, 99–101.
- [3] Nagaoka, S., Futamura, Y., Miwa, K., Awano, T., Yamauchi, K., Kanamaru, Y., Tadashi, K. and Kuwata, Y. (2001). Identification of novel hypocholesterolemic peptides derived from bovine milk β -lactoglobulin. *Biochem. Biophys. Res. Commun.* **281**, 11–17.
- [4] Matoba, N., Usui, H., Fujita, H. and Yoshikawa, M. (1999). A novel anti-hypertensive peptide derived from ovalbumin induces nitric oxide-mediated vasorelaxation in an isolated SHR mesenteric artery. *FEBS Lett.* **452**, 181–184.
- [5] Fujita, H. and Yoshikawa, M. (1999). LKPNM: a prodrug-type ACE-inhibitory peptide derived from fish protein. *Immunopharmacology* **44**, 123–127.
- [6] Yokoyama, K., Chiba, H. and Yoshikawa, M. (1992). Peptide inhibitors for angiotensin I-converting enzyme from thermolysin digest of dried bonito. *Biosci. Biotechnol. Biochem.* **56**, 1541–1545.
- [7] Shin, Z. I., Yu, R., Park, S. A., Chung, D. K., Ahn, C. W., Nam, H. S., Kim, K. S. and Lee, H. J. (2001). His-His-Leu, and angiotensin I converting enzyme inhibitory peptide derived from Korean soybean paste, exerts antihypertensive activity in vivo. *J. Agric. Food Chem.* **49**, 3004–3009.
- [8] Wu, J. and Ding, X. (2001). Hypotensive and physiological effect of angiotensin converting enzyme inhibitory peptides derived from soy protein on spontaneously hypertensive rats. *J. Agric. Food Chem.* **49**, 501–506.
- [9] Marczak, E. D., Usui, H., Fujita, H., Yang, Y., Yokoo, M., Lipkowski, A. W. and Yoshikawa, M. (2003). New antihypertensive peptides isolated from rapeseed. *Peptides* **24**, 791–798.
- [10] Yang, Y., Marczak, E. D., Yokoo, M., Usui, H. and Yoshikawa, M. (2003). Isolation and antihypertensive effect of angotensin I-converting enzyme (ACE) inhibitory peptides from spinach rubisco. *J. Agric. Food Chem.* **51**, 4897–4902.
- [11] Choi, H. S., Cho, H. Y., Yang, H. C., Ra, K. S. and Suh, H. J. (2001). Angiotensin-I converting enzyme inhibitor from *Grifola frondosa*. *Food Res. Int.* **34**, 177–182.
- [12] Lee, D. H., Kim, J. H., Park, J. S., Choi, Y. J. and Lee, J. S. (2004). Isolation and characterization of a novel angiotensin I-converting enzyme inhibitory peptide derived from the edible mushroom *Tricholoma giganteum*. *Peptides* **25**, 621–627.
- [13] Choi, S-K., Adachi, M. and Utsumi, S. (2004). Improved bile acid-binding ability of soybean glycinin A1a polypeptide by the introduction of a bile acid-binding peptide (VAWWMY). *Biosci. Biotechnol. Biochem.* **68**, 1980–1983.
- [14] Matoba, N., Doyama, N., Yamada, Y., Maruyama, N., Utsumi, S. and Yoshikawa, M. (2001). Design and production of genetically modified soybean protein with anti-hypertensive by incorporating potent analogue of ovokin(2-7). *FEBS Lett.* **497**, 50–54.
- [15] Onishi, K., Matoba, N., Yamada, Y., Doyama, N., Maruyama, N., Utsumi, S. and Yoshikawa, M. (2004). Optimal designing of β -conglycinin to genetically incorporate RPLKPW, a potent anti-hypertensive peptide. *Peptides* **25**, 37–43.
- [16] Takenaka, Y., Doyama, N., Maruyama, N., Utsumi, S. and Yoshikawa, M. (2004). Introduction of DPR, an enterostatin fragment peptide, into soybean β -conglycinin α' subunit by site-directed mutagenesis. *Biosci. Biotechnol. Biochem.* **68**, 253–256.

- [17] Takenaka, Y., Utsumi, S. and Yoshikawa, M. (2000). Introduction of enterostatin (VPDPR) and a related sequence into soybean proglycinin A1aB1b subunit by site-directed mutagenesis. *Biosci. Biotechnol. Biochem.* **64**, 2731–2733.
- [18] Takenaka, Y., Utsumi, S. and Yoshikawa, M. (2001). Introduction of a low molecular weight agonist peptide for complement C3a receptor into soybean proglycinin A1aB1b subunit by site-directed mutagenesis. *Biosci Biotechnol Biochem.* **65**, 1202–1205.
- [19] Dickinson, C. D., Hussein, E. H. A. and Nielsen, N. C. (1989). Role of posttranslational cleavage in glycinin assembly. *Plant Cell* **1**, 459–469.
- [20] Jung, R., Scott, M. P., Nam, Y. W., Beaman, T. W., Bassüner R., Saalbach I., Müntz K. and Nielsen N. C. (1998). The role of proteolysis in the processing and assembly of 11S seed globulins. *Plant Cell* **10**, 343–357.
- [21] Nam, Y. W., Jung, R. and Nielsen, N. C. (1997). Adenosine 5'-triphosphate is required for the assembly of 11S seed proglubulins in vitro. *Plant Physiol.* **115**, 1629–1639.
- [22] Katsube, T., Kurisaka, N., Ogawa, M., Maruyama, N., Ohtsuka, R., Utsumi, S. and Takaiwa, F. (1999). Accumulation of soybean glycinin and its assembly with the glutelins in rice. *Plant Physiol.* **120**, 1063–1073.
- [23] Takaiwa, F., Katsube, T., Kitagawa, S., Higasa, T., Koto, M. and Utsumi, S. (1995). High level accumulation of soybean glycinin in vacuole-derived protein bodies in the endosperm tissue of transgenic tobacco seed. *Plant Sci.* **111**, 39–49.
- [24] Utsumi, S., Kitagawa, S., Katsube, T., Kang, I. J., Gidamis, A. B., Takaiwa, F., Kito, M. (1993). Synthesis, processing and accumulation of modified glycinins of soybean in the seeds, leaves and stems of transgenic tobacco. *Plant Sci.* **92**, 191–202.
- [25] Choi, S-K., Adachi, M., Yoshikawa, M., Maruyama, N. and Utsumi, S. (2004). Soybean glycinin. A1aB1b subunit has a molecular chaperone-like function to assist the folding of the other subunit having low folding ability. *Biosci. Biotechnol. Biochem.* **68**, 300–333.
- [26] Iwami, K., Sakakibara, K., Ibuki, F. (1986). Involvement of post-digestion 'hydrophobic' peptides in plasma cholesterol-lowering effect of dietary plant proteins. *Agric. Biol. Chem.* **50**, 1217–1222.
- [27] Nagaoka, S., Miwa, K., Eto, M., Kuzuya, Y., Hori, G. and Yamamoto, K. (1999). Soy protein peptic hydrolysate with bound phospholipids decreases micellar solubility and cholesterol absorption in rats and Caco-2 cells. *J. Nutr.* **129**, 1725–1730.
- [28] Sugano, M., Goto, S., Yamada, Y., Yoshida, K., Hashimoto, Y., Matsuo, T. and Kimoto, M. (1990). Cholesterol-lowering activity of various undigested fractions of soybean protein in rats. *J. Nutr.* **120**, 977–985.
- [29] Sugano, M., Yamada, Y., Yoshida, K., Hashimoto, Y., Matsuo, T. and Kimoto, M. (1988). The hypocholesterolemic action of the undigested fraction of soybean protein in rats. *Atherosclerosis* **72**, 115–122.
- [30] Adachi, M., Takenaka, Y., Gidamis, A. B., Mikami, B. and Utsumi, S. (2001). Crystal structure of soybean proglycinin A1aB1b homotrimer. *J. Mol. Biol.* **305**, 291–305.
- [31] Gidamis, A. B., Mikami, B., Katsube, T., Utsumi, S. and Kito, M. (1994). Crystallization and preliminary X-ray analysis of soybean proglycinins modified by protein engineering. *Biosci. Biotechnol. Biochem.* **58**, 703–706.
- [32] Katsube, T., Gidamis, A. B., Kanamori, J., Kang, Il-J., Utsumi, S. and Kito, M. (1994). Modification tolerability of the hypervariable region of soybean proglycinin. *J. Agric. Food Chem.*; **42**, 2639–2645.
- [33] Sanger, F., Nicklen, S. and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl. Acad. Sci. USA* **74**, 5463–5467.

- [34] Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.
- [35] Kim, C. S., Kamiya, S., Kanamori, J., Utsumi, S. and Kito, M. (1990). High level expression, purification and functional properties of soybean proglycinin from *Escherichia coli*. *Agri. Biol. Chem.* **54**, 1543–1550.
- [36] Green, A. A. and Hughes, W. L. (1955). Methods in enzymology. Colowick SP, Kaplan NO, editors. New York: Academic Press, p. 67.
- [37] Maruyama, N., Katsube, T., Wada, Y., Oh M. H., Barba, de la Rosa A. P., Okuda, E., Nakagawa, S. and Utsumi, S. (1998). The roles of the N-linked glycans and extension regions of soybean β -conglycinin in folding, assembly and structural features. *Eur. J. Biochem.* **258**, 854–862.
- [38] Utsumi, S., Gidamis, A. B., Mikami, B. and Kito, M. (1993). Crystallization and preliminary x-ray crystallographic analysis of the soybean proglycinin expressed in *Escherichia coli*. *J. Mol. Biol.* **233**, 177–778.
- [39] Tandang, M. R. G., Atsuta, N., Maruyama, N., Adachi, M. and Utsumi, S. (2005). Evaluation of the solubility and emulsifying property of soybean proglycinin and rapeseed procruciferin in relation to structure modified by protein engineering. *J. Agric. Food Chem.* **53**, 8736–8744.
- [40] Tandang, M. R. G., Adachi, M., Inui, N., Maruyama, N. and Utsumi, S. (2004) Effects of protein engineering of canola procruciferin on its physicochemical and functional properties. *J. Agric. Food Chem.* **52**, 6810–6817.

Summary

In this thesis, the author studied on structure-physicochemical function relationships of soybean glycinins and proglycinins at subunit levels. Furthermore, the author improved physicochemical and physiological properties of soybean proglycinins by protein engineering. The findings in each chapter are summarized as follows:

Chapter I. Structure–Function Relationships of Soybean Glycinin at Subunit Levels Assessed Using Mutant Lines

Glycinin is a hexameric protein composed of five kinds of subunits. The subunits are classified into two groups, group I (A1aB1b, A1bB2, and A2B1a) and group II (A3B4 and A5A4B3). The author purified four mutant glycinins composed of only group I subunits (group I-glycinin), only group II subunits (group II-glycinin), only A3B4 (A3B4-glycinin), and only A5A4B3 (A5A4B3-glycinin) from mutant soybean lines. The physicochemical properties of these glycinin samples were compared with those of the normal glycinin (11S) composed of five kinds of subunits. The thermal stabilities (as measured by thermal denaturation midpoint temperatures) of 11S, group I-glycinin, and group II-glycinin were similar to each other, although that of A3B4-glycinin was significantly lower than those of the others. The orders of aromatic and aliphatic surface hydrophobicities were the same: A3B4-glycinin > group II-glycinin > A5A4B3-glycinin > 11S > group I-glycinin. The solubility of 11S as a function of pH at I = 0.5 was governed by that of group I-glycinin and followed this order at acidic pH: 11S = group I-glycinin > A3B4-glycinin > group II-glycinin = A5A4B3-glycinin. The order of emulsifying abilities was A5A4B3-glycinin > group II-glycinin > A3B4-glycinin > 11S > group I-glycinin. This order was consistent with the length of their hypervariable regions. Except for this relationship, there was no significant relationship among the other physicochemical properties of glycinins at subunit levels.

Chapter II. Structure–Function Relationships of Soybean Proglycinins at Subunit Levels

Although it is not possible to obtain individual molecular species of glycinin composed of only one kind of subunits from soybean seeds, it is possible to prepare such individual molecular species by employing the *E. coli* expression system. Even if they are in the form of proglycinin, a trimeric protein, the physicochemical properties of each individual subunit could be studied. cDNAs for individual subunits were cloned by RT-PCR method and expressed in *E. coli* using pET vector. The recombinant proglycinins were purified by ammonium sulfate fractionation and column chromatography in the form of homotrimers. Physicochemical properties such as molecular dimensions, solubility, surface hydrophobicity, thermal stability and emulsifying ability of individual proglycinins were studied. Molecular dimensions were proportional to molecular size for all proglycinins except A2B1a. Solubility was intrinsic to each proglycinin. At the ionic strength of 0.5, all proglycinins except A1aB1b showed very low solubility at acidic pH, but A1aB1b was soluble to higher than 60%. At ionic strength 0.08, all proglycinins exhibited isoelectric precipitation, although A2B1a and A1bB2 were not completely insoluble. The order of emulsifying ability ($A1bB2 < A2B1a < A5A4B3 < A3B4 \leq A1aB1b$) was not of the same for surface hydrophobicity ($A5A4B3 < A1aB1b \leq A3B4 < A1bB2 < A2B1a$) and thermal stability ($A1bB2 \ll A2B1a \leq A5A4B3 < A3B4 \leq A1aB1b$). These properties were different from those of the mature glycinins. Thus, the properties of mature glycinins are conferred by hexamer formation.

Chapter III. Protein Engineering of Soybean Proglycinin C-Terminus for Improving Physicochemical Properties

In the author's laboratory, it was found that introduction of the extension region (141 amino acids) of β -conglycinin α' subunit at the C-terminus of proglycinin A1aB1b resulted in improvement of emulsifying properties and a change in solubility profile. α' Extension region is high in charged residues, especially negatively charged residues. What structural characteristic, the length of the region, ratio of charged and hydrophilic residues and their distribution are important for the improvement? It is noted that the extension

region (125 amino acids) of β -conglycinin α subunit and the hypervariable region (103 amino acids) of A5A4B3 are also high in charged residues. Hydrophobicity profiles of α' and α extension and A5A4B3 hypervariable regions are different from each other. To investigate, the author introduced α' and α extension regions to A2B1a C-terminus, and α extension and A5A4B3 hypervariable regions to A1aB1b C-terminus. She also introduced an oligopeptide composed of twenty negatively or positively charged residues to the A1aB1b C-terminus.

All the mutants were expressed in *E. coli* and were purified. Their physicochemical properties such as a molecular dimension, thermal stability, surface hydrophobicity, solubility and emulsifying ability were studied. Analyses of a molecular dimension, thermal stability and surface hydrophobicity suggested that all mutants formed the proper conformation similar to that of the wild type. Solubility was intrinsic to each mutant as expected from the modification except for A1aB1bPos. It exhibited unexpected characteristics such as very poor solubility at ionic strength 0.5 at any pH. At ionic strength 0.08, the solubility of this mutant increased to around 70% when pH increased. These phenomena might be due to the combinations of the effects neutralization by electrostatic interaction of positively charged residues at the C-terminus with the negatively charged residues at the disordered regions II and IV, change of electrostatic potential of charged residues by pH, neutralization of charged residues by salt, and increase of hydrophobic interaction by salt. At ionic strength 0.5, the emulsifying abilities of all mutants were better than that of the wild type except A1aB1bPos and A1aB1bNeg, and at ionic strength 0.08 all mutants especially A1aB1bPos exhibited better emulsifying ability than that of the wild type, indicating that addition of oligo- and polypeptide rich in charged residues is generally a method suitable for improving emulsifying ability. The order of stability of the emulsion at both ionic strengths was $A1aB1b\alpha \geq A2B1a\alpha > A1aB1b\alpha' \geq A2B1a\alpha' \gg A1aB1bPos > A1aB1bA4IV \geq A1aB1bNeg > A2B1a > A1aB1b$. These results indicated that emulsion stability of proglycinin mutants depends on length, distribution of charged residues and hydrophilic/hydrophobic profile of the polypeptides introduced to the C-terminus of proglycinin.

Chapter IV. Design of Genetically Modified Soybean Proglycinin A1aB1b With Multiple Copies of Bioactive Peptide Sequences

Hypercholesterolemia is one of the most common lifestyle-related problems that causes obesity and diseases such as coronary heart disease and diabetes. A diet rich in foods with known hypocholesterolemic activity can potentially prevent these diseases. IIAEK derived from β -lactoglobulin is a peptide which has a hypocholesterolemic activity greater than that of a medicine, β -sitosterol. To create food proteins with hypocholesterolemic activity, the author introduced nucleotide sequences encoding tandem multimer of the IIAEK sequences into DNA regions corresponding to the five disordered regions of soybean glycinin A1aB1b subunit, and expressed the mutants in *Escherichia coli*. The expression level and solubility of the five mutants, each containing four IIAEK sequences in each disordered region, were compared. Overall, the expression level and solubility of the mutants with four IIAEK sequences at the disordered regions IV and V were the best followed by II > III > I. Further, introduction of the fifth IIAEK sequence to the disordered region IV did not decrease expression level and solubility. Increasing the number of IIAEK to seven and ten slightly decreased expression level, while the solubilities decreased to as low as 40 and 1%, respectively. Various mutations were combined to get a mutant containing as many IIAEK sequences as possible. Some of the resulting mutants were expressed in the soluble form. The mutant containing eight IIAEK from the combination of disordered regions IV and V showed the best balance of the expression level and solubility followed by the combination of disordered regions II and III. The soluble fractions of these mutants were purified by hydrophobic, gel filtration and ion-exchange column chromatography. Yields of IIAEK peptide released by in vitro digestion with trypsin from both mutants were around 80%. This is the first report that a large amount of the physiologically active peptide could be introduced into food protein.

Conclusions

1. Glycinin individual subunits had intrinsic physicochemical properties.
2. The emulsifying ability of glycinins showed correspondence with the length of their hypervariable regions. Except for this relationship, there was no significant relationship among the other physicochemical properties of glycinins at subunit level.
3. Thermal stability, surface hydrophobicity and emulsifying property of proglycinins were different from those of mature glycinins. Thus, the properties of mature glycinins are likely conferred by hexamer formation.
4. Emulsion stability of the proglycinins modified at the C-terminus region depended on the length and the distribution of charged residues of the introduced peptides.
5. Large amount of physiologically active peptide could be introduced into soybean glycinin.

Acknowledgements

The author wishes to express her utmost appreciation and gratitude to Dr. Shigeru Utsumi, Professor of Division of Agronomy and Horticultural Science, Graduate School of Agriculture, Kyoto University, for his inspiring guidance, untiring help and encouragement throughout her entire PhD work. The author is indebted to him for his meticulous attention, organization and his critical reading of her scientific papers and thesis drafts.

The author owes a debt of gratitude to Dr. Bunzo Mikami, associate professor of Division of Agronomy and Horticultural Science, Graduate School of Agriculture, Kyoto University and Dr. Evelyn Mae T. Mendoza, Professor of Institute of Plant Breeding, college of Agriculture, University of the Philippines, for their constructive, inspiration, generous help and encouragement throughout her research.

The author would like to express her gratitude to Dr. Nobuyuki Maruyama, Dr. Yukie Maruyama and Dr. Motoyasu Adachi, Graduate School of Agriculture, Kyoto University, for their constructive research work and valuable ideas, which always inspired her.

The author owes a great debt to Asian Youth Fellowship (AYF) program organized by Asia Science and Education for Economic Development (Asia SEED) and Japanese Government “Monbusho (Monbukagakusho) Scholarship” for financial support without which she would not be possible to pursue her graduate study at Kyoto University.

The author would like to express her sincere gratitude to Dr. Yasuki Matsumura, Professor of Division of Agronomy and Horticultural Science, Graduate School of Agriculture, Kyoto University; Dr. Masaaki Yoshikawa and Dr. Kohsaku Ohhinata, Professor and lecturer of division of Food Science and Biotechnology, Graduate School of Agriculture, Kyoto University; Dr. Willem F. Stevens and Dr. Sudip K. Rakshit, Professors of School of Environment, Resources and Development, Asian Institute of Technology (AIT), Thailand; Dr. Naofumi Kitabatake and Dr. Fumito Tani, Professor and

associate professor of Division of Food Science and Biotechnology, Graduate School of Agriculture, Kyoto University; Dr. Didier Montet and Dr. Daniel Pioch, Centre de Cooperation Internationale en Recherche Agronomique pour le Développement (CIRAD-AMIS), Montpellier, France; Mrs. Chikako Tani, Mr. Joy Bartolome A. Duldulao, Miss Takako Fukuda, Mr. Takashi Mori, Mr. Takayasu Motoyama, Mrs. Shiori Motoyama, Mr. Takafumi Itoh, Miss Kazuyo Nakatani and all members of Division of Agronomy and Horticultural Science, Graduate School of Agriculture, Kyoto University, for their kind collaborations, general helps, or encouragement.

The author owes a great dept of gratitude to her beloved mother: Neang Sok, brothers: Prak Sokhon, Prak Sith, Prak Sithorn, Prak Boran and Prak Boreth, sisters: Prak Phalla and Prak Bunna, Dr. Kriston Vizi Janos, and friends for their loves, supports, and encouragements throughout her life. This piece of work is dedicated to her loving father: Prak Toch and brothers: Prak Phalkun and Prak Sokhorn who are sadly missed and fondly remembered forever.

List of Publications

1. Maruyama, N., Park, K., Motoyama, S., Choi, S. K., Yagasaki, K., Ishimoto, M., Utsumi, S. (2004). Structure-physicochemical function relationships of soybean glycinin at subunit levels assessed by using mutant lines. *J. Agric. Food Chem.* **52**, 8197–8201.
2. Prak, K., Nakatani, K., Katsube, T. T., Adachi, M., Maruyama, N., Utsumi, S. (2005). Structure–function relationships of soybean proglycinins at subunit levels. *J. Agric. Food Chem.* **53**, 3650–3657.
3. Prak, K., Maruyama, Y., Maruyama, N., Utsumi, S. (2006). Design of genetically modified soybean proglycinin A1aB1b with multiple copies of bioactive peptide sequences. *Peptides* **27**, 1179–1186.
4. Prak, K., Nakatani, M., Maruyama, N., Utsumi, S. (2007). C-terminus engineering of soybean proglycinin: improvement of emulsifying properties. *Protein Eng. Des. Sel.* **20**, 433–442.